JCSM
Journal of Clinical Sleep Medicine

SPECIAL ARTICLES

Clinical Practice Guideline for the Pharmacologic Treatment of Chronic Insomnia in Adults: An American Academy of Sleep Medicine Clinical Practice Guideline

Michael J. Sateia, MD[1]; Daniel J. Buysse, MD[2]; Andrew D. Krystal, MD, MS[3]; David N. Neubauer, MD[4]; Jonathan L. Heald, MA[5]

[1]Geisel School of Medicine at Dartmouth, Hanover, NH; [2]University of Pittsburgh School of Medicine, Pittsburgh, PA; [3]University of California, San Francisco, San Francisco, CA; [4]Johns Hopkins University School of Medicine, Baltimore, MD; [5]American Academy of Sleep Medicine, Darien, IL

Introduction: The purpose of this guideline is to establish clinical practice recommendations for the pharmacologic treatment of chronic insomnia in adults, when such treatment is clinically indicated. Unlike previous meta-analyses, which focused on broad classes of drugs, this guideline focuses on individual drugs commonly used to treat insomnia. It includes drugs that are FDA-approved for the treatment of insomnia, as well as several drugs commonly used to treat insomnia without an FDA indication for this condition. This guideline should be used in conjunction with other AASM guidelines on the evaluation and treatment of chronic insomnia in adults.

Methods: The American Academy of Sleep Medicine commissioned a task force of four experts in sleep medicine. A systematic review was conducted to identify randomized controlled trials, and the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) process was used to assess the evidence. The task force developed recommendations and assigned strengths based on the quality of evidence, the balance of benefits and harms, and patient values and preferences. Literature reviews are provided for those pharmacologic agents for which sufficient evidence was available to establish recommendations. The AASM Board of Directors approved the final recommendations.

Recommendations: The following recommendations are intended as a guideline for clinicians in choosing a specific pharmacological agent for treatment of chronic insomnia in adults, when such treatment is indicated. Under GRADE, a STRONG recommendation is one that clinicians should, under most circumstances, follow. A WEAK recommendation reflects a lower degree of certainty in the outcome and appropriateness of the patient-care strategy for all patients, but should not be construed as an indication of ineffectiveness. GRADE recommendation strengths do not refer to the magnitude of treatment effects in a particular patient, but rather, to the strength of evidence in published data. Downgrading the quality of evidence for these treatments is predictable in GRADE, due to the funding source for most pharmacological clinical trials and the attendant risk of publication bias; the relatively small number of eligible trials for each individual agent; and the observed heterogeneity in the data. The ultimate judgment regarding propriety of any specific care must be made by the clinician in light of the individual circumstances presented by the patient, available diagnostic tools, accessible treatment options, and resources.

1. We suggest that clinicians use suvorexant as a treatment for sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
2. We suggest that clinicians use eszopiclone as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
3. We suggest that clinicians use zaleplon as a treatment for sleep onset insomnia (versus no treatment) in adults. (WEAK)
4. We suggest that clinicians use zolpidem as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
5. We suggest that clinicians use triazolam as a treatment for sleep onset insomnia (versus no treatment) in adults. (WEAK)
6. We suggest that clinicians use temazepam as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
7. We suggest that clinicians use ramelteon as a treatment for sleep onset insomnia (versus no treatment) in adults. (WEAK)
8. We suggest that clinicians use doxepin as a treatment for sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
9. We suggest that clinicians not use trazodone as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
10. We suggest that clinicians not use tiagabine as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
11. We suggest that clinicians not use diphenhydramine as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
12. We suggest that clinicians not use melatonin as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
13. We suggest that clinicians not use tryptophan as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. (WEAK)
14. We suggest that clinicians not use valerian as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. (WEAK)

Keywords: insomnia, treatment, pharmacologic, guideline

Citation: Sateia MJ, Buysse DJ, Krystal AD, Neubauer DN, Heald JL. Clinical practice guideline for the pharmacologic treatment of chronic insomnia in adults: an American Academy of Sleep Medicine clinical practice guideline. *J Clin Sleep Med.* 2017;13(2):307–349.

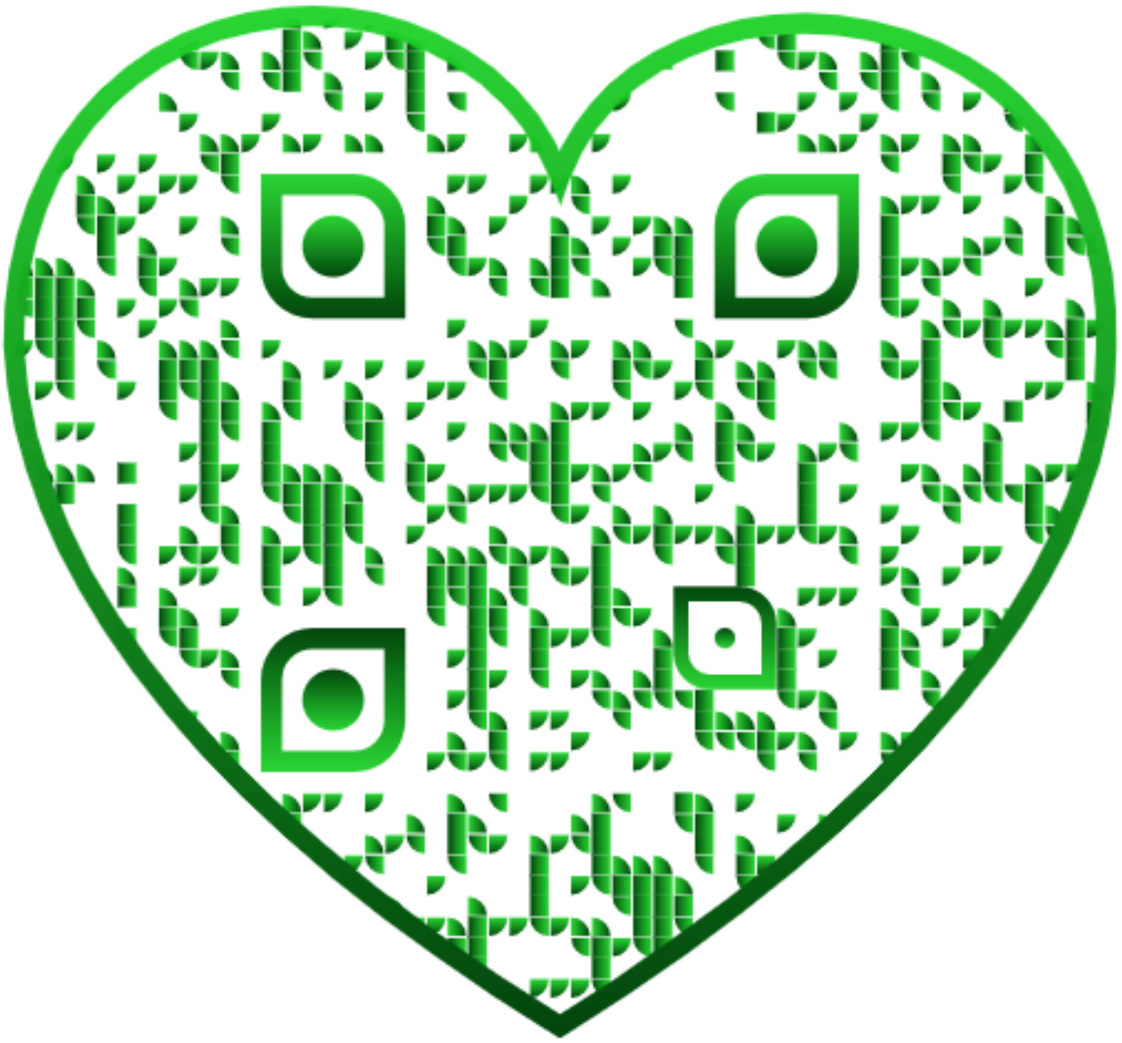


TABLE OF CONTENTS

Introduction *308*
Background *309*
Methodology *312*
Clinical Practice Recommendations *315*
Orexin receptor agonists
Suvorexant *317*
Benzodiazepine receptor agonists
Eszopiclone *318*
Zaleplon *321*
Zolpidem *323*
Benzodiazepines
Triazolam *326*
Temazepam *327*
Melatonin agonists
Ramelteon *329*
Heterocyclics
Doxepin *331*
Trazodone *332*
Anticonvulsants
Tiagabine *333*
Over-the-counter preparations
Diphenhydramine *334*
Melatonin *335*
L-tryptophan *337*
Valerian *338*
Literature Reviews *338*
Estazolam *338*
Quazepam *339*
Flurazepam *340*
Oxazepam *341*
Quetiapine *341*
Gabapentin *341*
Paroxetine *341*
Trimipramine *342*
Discussion and Future Directions *342*

INTRODUCTION

Aims

This clinical practice guideline was initiated at the request of the Board of Directors of the American Academy of Sleep Medicine (AASM), who also reviewed this document and provided feedback. No formal clinical practice guidelines for the pharmacological treatment of insomnia have previously been issued by the AASM, despite the fact that this remains, by far, the most common approach to therapy, after treatment of comorbidities. Pharmacotherapy is one of two major approaches to treatment, the alternative being cognitive behavioral therapies for insomnia (CBT-I), already identified as a standard of treatment. This paper does not directly address the relative benefits of these two approaches. Rather, the conclusions and recommendations regarding pharmacotherapy must be considered within the context of specific treatment goals, comorbidities, prior treatment responses, availability, safety, patient preference and cost considerations. Despite the clearly favorable benefit to risk ratio of CBT-I, not all patients with an insomnia disorder can and will derive benefit from this treatment alone. This failure may result from inability to access such treatment (due to availability, cost restraints, etc.), inability or unwillingness to participate in the therapy, or treatment non-responsiveness. Thus, pharmacotherapy, alone or in combination with CBT-I, must continue to be considered a part of the therapeutic armamentarium, as it currently is for perhaps 25% of the population.[1] Unfortunately, many individuals use medications or substances (e.g. over-the-counter sleep aids or alcohol) which are not demonstrated to be effective in managing insomnia and/or have significant potential for harm. For the estimated 3.5% to 7% of individuals receiving prescription medication for sleep disturbance,[2–4] significant knowledge gaps and anxieties about the proper usage of these agents exists among the prescribers.

This paper includes a systematic review and meta-analyses which provides the basis of the initial AASM clinical practice guideline for pharmacological management of insomnia. The aims of the present analysis are: (1) to determine the efficacy of individual prescription and non-prescription medications for treatment of insomnia; (2) to assess the efficacy of individual medications for specific sleep complaints (i.e. difficulty initiating sleep/difficulty maintaining sleep); (3) to evaluate the potential for adverse effects of these drugs; (4) to consider the evidence concerning efficacy and adverse effects in delineating evidence-based guidelines for the use of pharmacotherapy in the management of chronic insomnia; and (5) to offer recommendations for optimizing quality and uniformity of future investigations.

This clinical practice guideline is intended to serve as one component in an ongoing assessment of the individual patient with insomnia. As discussed elsewhere,[5–7] a comprehensive initial evaluation should include a detailed history of sleep complaints, medical and psychiatric history, and medication/substance use. These factors, together with patient preferences and treatment availability, should be used to select specific treatments for specific patients. This clinical practice guideline is not intended to help clinicians determine which patient is appropriate for pharmacotherapy. Rather, it is intended to provide recommendations regarding specific insomnia drugs once the decision has been made to use pharmacotherapy. This guideline is also not intended to recommend one drug over another. Very few comparative efficacy studies have been conducted among these agents. Rather, the guideline provides a recommendation and evidence base for each individual drug. The selection of a particular drug should rest on the evidence summarized here, as well as additional patient-level factors, such as the optimal pharmacokinetic profile, assessments of benefits versus harms, and past treatment history.

This guideline should be used in conjunction with other AASM guidelines on the evaluation and treatment of chronic insomnia. These guidelines indicate that CBT-I is a standard of treatment and that such treatment carries a significantly favorable benefit:risk ratio. Therefore, based on these guidelines, all patients with chronic insomnia should receive CBT-I as a primary intervention. Medications for *chronic* insomnia disorder should be considered mainly in patients who are unable

to participate in CBT-I, who still have symptoms despite participation in such treatments, or, in select cases, as a temporary adjunct to CBT-I.

Clinical Guidelines and Practice Parameters

The AASM has issued several guidelines, reviews, and practice parameters related to the assessment and management of insomnia. A 2000 review and practice parameter paper addressed the comprehensive evaluation of chronic insomnia.[5,6] Non-pharmacological management of insomnia has been the subject of two practice parameter papers.[8–11] No formal, evidence-based standards of practice for pharmacological treatment of insomnia have been published, although clinical guidelines addressing this topic have been issued by various groups. The Standards of Practice Committee of the AASM addressed non-prescription treatments for insomnia in a 2006 paper[12] which concluded that there is sparse or little evidence to support use of these agents for insomnia. Preliminary but conflicting evidence was noted for valerian and first-generation H_1 antagonists for short-term use. A 2005 National Institutes of Health consensus conference[13] on manifestations and management of chronic insomnia found moderate-to-high-grade evidence to support the efficacy of both cognitive-behavioral therapies and benzodiazepine agonists in the short-term management of insomnia, but noted a relative paucity of data concerning long-term usage of such treatments, despite the chronicity of the condition. Little evidence supporting efficacy of other widely used treatments (sedating antidepressants and non-prescription agents) was found.

A 2008 AASM clinical guideline paper on the evaluation and management of chronic insomnia defined psychological and behavioral therapies as a standard of treatment (the highest level of recommendation at that time).[7] No specific level of recommendation was offered for pharmacological therapies, but the consensus recommendation was that such treatment, when used, should be accompanied by cognitive-behavioral therapies whenever possible. Short/intermediate acting benzodiazepine receptor agonists (benzodiazepines [BZDs] or newer BZD receptor agonistic modulators [BzRAs]) or ramelteon were recommended as first-line pharmacotherapy. Other drugs, such as sedating antidepressants or anticonvulsant medications were recommended as second- or third-line agents, particularly when comorbidities (e.g. mood disorder or epilepsy) are present. Other, non-prescription drugs such as over-the counter antihistamine sleeping aids and herbal/nutritional agents were not recommended due to lack of demonstrated efficacy as well as safety concerns.

A consensus statement from the British Association for Psychopharmacology[14] assessed evidence related to chronic insomnia, including management issues, and came to similar conclusions. CBT interventions were recommended as first-line treatment. BzRAs were found effective for short-term use, although degradation of improvement following discontinuation of hypnotic was noted to be of concern. Limited evidence and toxicity concerns were cited for other prescription and non-prescription agents, although prolonged-release melatonin was recommended as a first-line treatment for insomnia in persons over 55 years.

In May 2016, the American College of Physicians published its own clinical practice guideline for the management of chronic insomnia.[15] This guideline makes two major recommendations. The first is that all patients with chronic insomnia receive CBT-I as the initial treatment intervention. This is a strong recommendation based on moderate quality evidence. The second is that a shared decision-making approach be employed by clinicians in determining whether pharmacotherapy should be employed for those patients who did not achieve adequate response with CBT-I (weak recommendation based on low quality evidence). The guideline notes that there was insufficient evidence to draw conclusions regarding the overall efficacy of pharmacotherapy in the insomnia population. More specifically, there was also insufficient evidence to determine the efficacy of benzodiazepines, trazodone and melatonin in the management of chronic insomnia. Studies of more recent generation sleep aids such as BzRAs, doxepin and suvorexant found improvement in a number of sleep outcome variable but, as is the case with our own guideline, much of the evidence was of low quality. Although evidence is presented for individual drugs, there were no specific recommendations made for single agents. Finally, there was insufficient evidence found to determine the balance of benefits versus harms.

BACKGROUND

Insomnia disorder is defined in the International Classification of Sleep Disorders, Third Edition[16] as a complaint of trouble initiating or maintaining sleep which is associated with daytime consequences and is not attributable to environmental circumstances or inadequate opportunity to sleep. The disorder is identified as chronic when it has persisted for at least three months at a frequency of at least three times per week. When the disorder meets the symptom criteria but has persisted for less than three months, it is considered short-term insomnia.

Occasional, short-term insomnia affects 30% to 50% of the population.[17] The prevalence of chronic insomnia disorder in industrialized nations is estimated to be at least 5% to 10%.[18,19] In medically and psychiatrically ill populations, as well as in older age groups, the prevalence is significantly higher. Chronic insomnia is associated with numerous adverse effects on function, health, and quality of life. Epidemiologic studies demonstrate marked impairment in functional status among those with chronic insomnia.[20,21] Increased rates of work absenteeism,[22] and occupational and motor vehicle accidents have also been widely reported.[23,24] Persistent insomnia has been identified in multiple studies as a significant risk factor for the development of psychiatric disorders, especially mood disorder.[25,26] This condition is also associated with increased risk of relapse for depression and alcoholism, as well as adverse effects in chronic pain populations. More recent investigations suggest that chronic insomnia is associated with increased risk for cardiovascular disease. In particular, insomnia with objectively short sleep time is a significant risk factor for the development of hypertension.[27,28]

Chronic insomnia imposes substantial economic burdens on society.[29–31] Estimation of the direct and indirect costs

of chronic insomnia are complicated by many confounding variables, but virtually all analyses of these costs indicate substantially higher economic burden for an insomnia population. Direct cost analysis demonstrates significantly higher utilization of emergency and office health care visits as well as greater cost for prescription drugs.[32] Likewise, indirect costs in the form of work absenteeism, loss of productivity, and insomnia-related accidents contribute significantly to the economic burden of the disorder. In the United States, a 2009 study[33] found that direct and indirect costs for insomnia patients were in excess of $2,000/year greater than those of a matched non-insomnia group. Total direct and indirect cost estimates for insomnia in the United States differ substantially due to variability in methodologies. Nevertheless, estimates suggest direct costs of $2–16 billion per year and indirect costs of $75–100 billion. The latter are accounted for in large part by worker absenteeism, presenteeism (lower productivity due to daytime impairment), and work-related accidents.[29]

General treatment measures for insomnia include the treatment of comorbid medical and psychiatric conditions, modifying sleep-interfering medications and substances, and optimizing the sleep environment. Specific treatments for insomnia fall into two primary categories. Non-pharmacological therapies, largely cognitive behavioral in nature, have been the subject of numerous meta-analyses and practice guidelines.[10,34–37] Pharmacological therapy, including over-the-counter sleep aids and alcohol, is the most widely used treatment for insomnia, yet no evidence-based clinical practice guidelines have been published to date by the AASM. This paper includes a systematic review and meta-analyses which provide the basis of the initial AASM clinical practice guideline for pharmacological management of insomnia.

History of Hypnotic Usage

Pharmacological agents have been used for the treatment of insomnia throughout much of recorded history. Prior to the 20th century, opioids, various herbal preparations, bromide salts, and alcohol were the primary hypnotic alternatives. Through the first half of the 20th century, barbiturate and related compounds became the most commonly used agents for management of anxiety and sleep disturbance, as well as epilepsy. By mid-century, however, the adverse side effects and lethal overdose potential of these agents became recognized, contributing to curtailment of use.

The first BZD, chlordiazepoxide, was introduced to the United States market in 1963, followed shortly by diazepam. Flurazepam, the first benzodiazepine approved by the Food and Drug Administration (FDA) as a hypnotic, became available in 1970 and rapidly supplanted the use of barbiturates and similar compounds for treatment of insomnia. Zolpidem, the first United States nonbenzodiazepine, benzodiazepine receptor agonist (non-BZD, or BzRA) hypnotic, became available in 1992 and remains the most widely prescribed hypnotic medication, accounting for 87.5% of all BzRA prescriptions in a recent survey of hypnotic use.[38] Since 2005, a melatonin agonist (ramelteon), a low dose form of the sedating tricyclic medication (doxepin), and, most recently, an orexin receptor antagonist (suvorexant) have entered the United States market.

Current Hypnotic Usage

Hypnotic prescribing practices have varied in recent decades as availability of various agents and safety concerns have evolved. Despite the development of numerous BZD hypnotic medications of varying durations of action, the overall frequency of hypnotic prescriptions for insomnia declined during the two decades from 1970–1990, from 3.5% to 2.5%.[39] Due to apparent concerns regarding the potential for tolerance and dependency with BZD use, physicians increasingly prescribed sedating antidepressants "off label," especially trazodone, despite the absence of efficacy studies for this or any other sedating antidepressants for treatment of insomnia. Survey of office-based physician prescribing practices for the period 1987–1996 revealed an over 50% decline in BZD hypnotic prescriptions accompanied by a nearly 150% increase in trazodone prescriptions.[40] Overall prescriptions for insomnia declined by about 25% during this period. A more recent study,[38] utilizing the National Health and Nutrition Examination Survey (NHANES) data from 1999–2010, analyzed the frequency of usage of medications commonly used for insomnia. This includes BZDs approved for treatment of insomnia, BzRAs, ramelteon, trazodone, doxepin and quetiapine. The authors report that just under 3% of the sample population used one of these agents within the past month. In contrast to the apparent trends of preceding decades, frequency of usage of any medication commonly used for insomnia increased over the decade, from 2.0% in the first year sampled to 3.5% in the final year (2009–2010). BzRAs, predominantly zolpidem, were most commonly prescribed (1.23% of the population), followed by trazodone (0.97%), BZDs (0.4%), quetiapine (0.32%) and doxepin. However, it should be noted in this and other studies that other agents—especially BZDs not approved for insomnia, other antidepressants, antipsychotics, and analgesics—are not included in these data. It seems likely that the true prevalence of medication use for sleep disturbance is higher than these figures suggest. In fact, a subsample analysis of the NHANES data from 2005–2008 found that approximately 19% of respondents reported use of at least one pill or medication for sleep in the past month. The 2005 National Sleep Foundation's (NSF) survey of sleep habits found that 7% of respondents report using a prescription medication to improve sleep at least a few nights per month.[41]

Physicians and other health care providers have consistently expressed reservations about the use of medication, particularly BZDs and BzRAs, to treat insomnia. They cite concerns regarding safety and dependency as key issues. However, they also note a lack of awareness and/or availability of alternative treatments.[42] Many favor an initial approach of treating associated comorbidities and advising good sleep hygiene.[43] An ever-increasing amount of data makes it clear that the latter approach is very often unsuccessful, leaving providers feeling compelled to prescribe medications. Most of those surveyed recognize the need for additional, non-pharmacological treatment for their patients, but cite a number of barriers to acquiring such treatment.[44]

Data concerning use of non-prescription agents for sleep promotion are limited. The aforementioned NSF survey reported that nearly one in four respondents used some form of sleep aid "at least several times per month." Eleven percent stated that they used alcohol to help sleep, 9% used over-the counter sleep aids, and 2% used melatonin.

Previous Meta-Analyses

Several meta-analyses of pharmacotherapy for insomnia have been conducted. Nowell and colleagues[45] conducted a meta-analysis of 22 randomized controlled trials (RCTs) of BZDs and zolpidem for treatment of primary insomnia published from 1966 to 1996. They found moderate effect sizes (d = 0.56 to 0.71 for key sleep variables) for improvement with these agents, but noted limitations in the duration of trials and lack of follow-up study regarding outcome. A 2000 study commissioned by the Canadian Medical Association[46] evaluated 45 RCTs (n = 2,672) of BZDs for treatment of primary insomnia. This investigation found reduction in sleep latency (non-significant in objective [polysomnography; PSG] assessment but significant in subjective reports) and a somewhat larger and significant increase in total sleep time by both objective and subjective reports. The authors also note an increase in adverse events with BZDs (pooled odds ratio for any adverse event = 1.8) and call into question the risk/benefit ratio for these agents.

A comparative evaluation of the efficacy of hypnotic drugs was conducted by the National Centre for Clinical Excellence of the UK.[47] In summary, the analysis found little difference among the numerous BZDs and BzRAs among the 24 studies which directly compared more than one drug. Some small differences were noted, such as shorter sleep latency but less total sleep time with zaleplon when compared to zolpidem. On the whole, major differences in adverse effects were not observed between drugs. Meta-analyses in this report were few due to limitations of data reporting.

Glass and colleagues[48] compared benefits versus risks for all sedative hypnotic agents in a meta-analysis of RCTs of active agent versus placebo or other active compound in populations > 60 years of age and free of contributing comorbidities. They reported a small effect size for improvement in sleep quality (d = 0.14). Separate analysis of BZDs alone yielded a somewhat more robust improvement in quality (d = 0.37). Significant but modest increase in total sleep time (TST) and reduction in number of awakenings (NOA) was also found for all sedative-hypnotics and for the BZD group alone, although effect sizes are not reported for these variables. Cognitive side effects were more common with sedative-hypnotics. The authors note that, with respect to the sleep quality measures reported for all sedative hypnotics, the number needed to treat is 13, while the number needed to harm is 6, thereby indicating an unfavorable risk/benefit ratio for this population. Independent analysis of this ratio for BZDs alone was not conducted.

A 2007 meta-analysis[49] evaluated 105 RCTs of BZDs, BzRAs and antidepressant medications for treatment of chronic insomnia in the adult populations regardless of comorbidities. In summary, the analysis indicates moderate and significant improvement in major sleep parameters with both BZDs and BzRAs in both objective (PSG) and subjective (sleep diary) assessments, with the exception of PSG results for wake after sleep onset (WASO) and TST, which yielded results just below the range of significance. Far fewer studies were available for antidepressants. These showed significant reduction in sleep latency and a non-significant trend toward reduced WASO. Four studies utilizing PSG measures showed substantial improvement in TST (79.6 min) while single subjective data set suggested reduction in TST compared to placebo. The authors note substantial heterogeneity of data which was reduced in subgroup analyses by type of drug. Between-groups comparisons showed no significant efficacy differences between BZDs and non-BZDs. All three groups demonstrated significantly higher rates of adverse events versus placebo. BZDs exhibited a higher rate of adverse events than BzRAs.

Huedo-Medina and colleagues[50] conducted systematic review and meta-analysis of data on BzRAs submitted to the United States Food and Drug Administration from 15 studies. They found that BzRAs produce significant reduction of sleep latency by both objective and subjective measures with effect sizes of 0.36 and 0.33, respectively. Other sleep variables did not show significant differences but limited data reporting on these variables precluded definitive conclusions.

Winkler and Doering[51] analyzed data from 31 randomized controlled trials of sleep-promoting substances used for treatment of primary insomnia. Studies included BZDs, BzRAs, melatonin agonists, antidepressants and other sedating compounds. Only studies which included objective (PSG) data were considered. The meta-analysis revealed that both BZDs and BzRAs were significantly more effective than antidepressants. Both demonstrated small to moderate effect sizes for major sleep variables. BZDs were somewhat superior to BzRAs for subjective sleep latency (SL). No analysis of treatment-emergent adverse events was reported.

Finally, Wilt and colleagues[52] conducted a systematic review and meta-analyses of 35 randomized, controlled trials of at least 4 weeks duration, and harms information from 11 long-term observational trials. Their review found that eszopiclone, zolpidem, and suvorexant improved short-term outcomes, with small effect sizes and low-to-moderate strength evidence. They also found that evidence for BZDs, melatonin agonists, and antidepressants was insufficient or of too-low strength. Finally, they concluded that there is insufficient evidence to determine the comparative effectiveness or long-term efficacy of pharmacotherapies for insomnia.

In summary, these meta-analyses suggest small to moderate effect sizes for most major sleep outcome variables with both BZDs and BzRAs. However, some of these analyses report significant increases in treatment-emergent adverse events and raise concerns regarding their relative risk:benefit ratio. Data supporting the use of sedating antidepressants in the treatment of insomnia are scant. Overall, the studies are limited by lack of availability of high quality evidence and considerable variability in design and methodology across investigations. All of these analyses addressed efficacy only for major drug groups (e.g., BZDs, BzRAs), failing to address issues of safety or efficacy for individual agents.

Table 1—PICO question parameters.

Population	Intervention	Comparison	Outcomes
Adult patients diagnosed with primary chronic insomnia	1. Diphenhydramine † 2. Doxepin* 3. Eszopiclone* 4. Melatonin † 5. Ramelteon* 6. Suvorexant* 7. Temazepam* 8. Tiagabine** 9. Trazodone** 10. Triazolam* 11. Tryptophan † 12. Valerian ± hops † 13. Zaleplon* 14. Zolpidem*	Placebo control	Sleep latency (SL) Total sleep time (TST) Wake after sleep onset (WASO) Quality of sleep (QOS) Sleep efficiency (SE) Number of awakenings (NOA)

* = FDA-approved, indicated for the treatment of insomnia. ** = FDA-approved, off-label usage for the treatment of insomnia. † = Over-the-counter medication. Sleep latency is defined as the time to fall asleep following bedtime. PSG sleep latency may be reported as time to onset of first epoch of N1 (Stage 1) sleep, or, in more recent studies, as latency to persistent sleep (LPS), or time to onset of first 10 consecutive min of sleep. Total sleep time is defined as the total time spent in bed, minus sleep latency and wake after sleep onset. Wake after sleep onset is defined as the sum of wake times from sleep onset to the final awakening. Quality of sleep is a patient-reported measure, the definition of which varies by measurement tools and patient perceptions. Sleep efficiency is defined as the percentage of time spent in bed during which sleep occurs; it is calculated as (TST / time in bed) × 100. Number of awakenings is defined as the number of awakenings after sleep onset, excluding the final awakening.

Table 2—"Critical" outcomes by intervention.

	TST	SL	WASO	QOS
Diphenhydramine	✓	✓	✓	✓
Doxepin	✓	✓	✓	✓
Eszopiclone	✓	✓	✓	✓
Melatonin		✓		✓
Ramelteon		✓		✓
Suvorexant	✓	✓	✓	✓
Temazepam	✓	✓	✓	✓
Tiagabine	✓		✓	✓
Trazodone	✓	✓	✓	✓
Triazolam		✓		✓
Tryptophan		✓	✓	✓
Valerian-hops		✓		✓
Zaleplon	✓	✓		✓
Zolpidem	✓	✓	✓	✓

TST = total sleep time, SL = sleep latency, WASO = wake after sleep onset, QOS = quality of sleep.

METHODOLOGY

Expert Task Force

In order to develop this clinical practice guideline, the AASM commissioned a task force composed of content experts in the field of insomnia, an AASM Board of Directors liaison, and AASM Science and Research Department staff members. Prior to appointment, the content experts were required to disclose all potential conflicts of interest according to the AASM's policy. In accordance with the AASM's conflicts of interest policy, task force members with a Level 1 conflict were not allowed to participate. Task force members with a Level 2 conflict were required to recuse themselves from any related discussion or writing responsibilities. All relevant conflicts of interest are listed in the Disclosures section.

PICO Questions

A PICO (Patient, Population or Problem, Intervention, Comparison, and Outcomes) question template was developed to be the focus of this guideline:

> "In adult patients diagnosed with primary chronic insomnia, how does [intervention] improve [outcomes], compared to placebo?"

The PICO question template was approved by the AASM Board of Directors. The task force identified the pharmacological interventions of interest, based on FDA approval status and common off-label usage. Based on their expertise, the task force developed a list of patient-oriented clinically relevant outcomes that are indicative of whether a treatment should be recommended for clinical practice. The task force then rated their relative importance and selected the top six outcomes. The following outcomes were determined to be "critical" or "important" for clinical decision making across all interventions: sleep latency, wake after sleep onset, total sleep time, quality of sleep, number of awakenings, and sleep efficiency (**Table 1**). The task force then determined which outcomes were "critical" for clinical decision making for each individual intervention (**Table 2**). Lastly, clinical significance thresholds, used to determine if a change in an outcome was clinically significant, were defined for each outcome by task force clinical judgement, prior to statistical analysis (**Table 3**). These decisions were made by nominal consensus of the task force, based

Table 3—Clinical significance threshold.

Outcome	Measurement Tool [a]		
	Polysomnography	Actigraphy	Subjective
Sleep latency (SL), min	10	10	20
Total sleep time (TST), min	20	20	30
Wake after sleep onset (WASO), min	20	20	30
Quality of sleep (QOS), varies [b]	Varies	Varies	Varies
Sleep efficiency (SE), %	5	5	10
Number of awakenings (NOA), n	2	2	0.5

[a] Clinical significance was judged to be present when a specific agent led to a mean change in the outcome of this magnitude, compared to placebo.
[b] For standardized mean difference (SMD), an effect size of 0.5 is considered clinically significance (based on Cohen's d).

on their expertise and familiarity with the literature and clinical practice.

Literature Searches, Evidence Review and Data Extraction

Multiple literature searches were performed by the AASM research staff using the PubMed database throughout the guideline development process (see **Figure 1**). Keywords and Medical Subject Headings (MeSH) terms were:

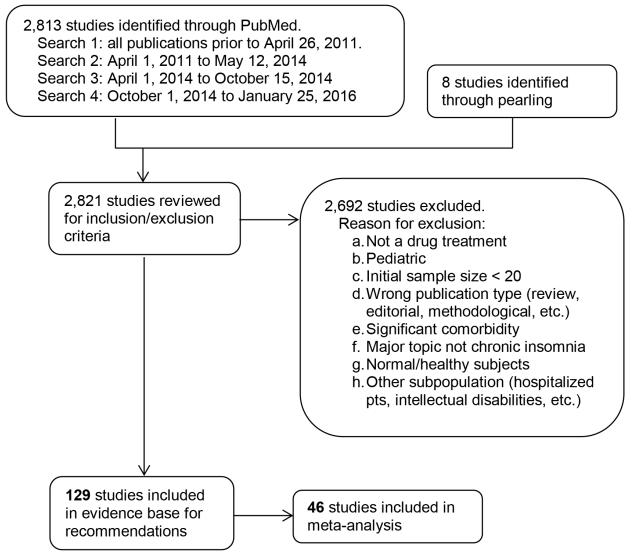
- insomnia OR sleep initiation and maintenance disorder NOT transient AND
- clinical trial OR randomized controlled trial
- NOT editorial, letter, comment, case reports, biography, review

The full literature search string can be found in the supplemental material. Searches were performed on April 26, 2011 (search 1), May 12, 2014 (search 2), October 15, 2014 (search 3), and January 25, 2016 (search 4). Based on their expertise and familiarity with the insomnia literature, task force members submitted additional relevant literature and screened reference lists to identify articles of potential interest. This served as a "spot check" for the literature searches to ensure that important articles were not missed.

Abstracts from all retrieved articles were individually assessed by two task force members to determine whether the publication should be included or excluded from further consideration in the project. Exclusion criteria can be found in **Figure 1**. A total of 129 publications were approved for inclusion.

Full texts of accepted articles were reviewed and data pertaining to GRADE[53] for the outcomes of interest were extracted into spreadsheets by AASM staff. All data pertaining to adverse events were extracted into separate spreadsheets. Articles that met inclusion criteria but did not report outcomes of interest were rejected from the final evidence base. If outcome data were not presented in the format necessary for statistical analysis (i.e., mean, standard deviation, and sample size), the authors were contacted in an attempt to obtain the necessary data. Finally, clinicaltrials.gov was used as a final resource for attempting to obtain data necessary for completing statistical analyses. If the necessary data were not available from the publication, the author, or clinicaltrials.gov, the paper was included in the evidence base as supporting evidence, but was not used for statistical analysis or for determining quality of evidence. Of the 129 accepted publications, 46 were included in the statistical and meta-analysis.

Figure 1—Evidence base flow diagram.


2,813 studies identified through PubMed.
Search 1: all publications prior to April 26, 2011.
Search 2: April 1, 2011 to May 12, 2014
Search 3: April 1, 2014 to October 15, 2014
Search 4: October 1, 2014 to January 25, 2016
8 studies identified through pearling
2,821 studies reviewed for inclusion/exclusion criteria
2,692 studies excluded.
Reason for exclusion:
a. Not a drug treatment
b. Pediatric
c. Initial sample size < 20
d. Wrong publication type (review, editorial, methodological, etc.)
e. Significant comorbidity
f. Major topic not chronic insomnia
g. Normal/healthy subjects
h. Other subpopulation (hospitalized pts, intellectual disabilities, etc.)
129 studies included in evidence base for recommendations
46 studies included in meta-analysis

For some drugs, none of the accepted publications provided data that could be used for statistical analysis. In these cases, the task force did not make a recommendation, but provided a literature review of these accepted papers instead. These publications are not included in **Figure 1**.

Statistical and Meta-Analysis

For outcomes of interest, data from baseline and last-treatment time points were used for all statistical and meta-analyses. Data from crossover trials were treated as parallel groups. Change-from-baseline values were also used for statistical and meta-analyses, when the change-from-baseline standard deviation was provided or could be calculated from the provided statistic. Standardized mean difference (SMD) was used for meta-analyses of quality of sleep (QOS) when data were reported using variable scales. Analyses were limited to FDA-approved doses. For adverse events, all data presented in the accepted papers were used for statistical and meta-analysis. All calculations and meta-analyses were performed using Review Manager 5.3 software. Whenever possible, meta-analyses were

performed by pooling data across studies for each outcome and adverse event. The evidence was grouped for analysis based on the drug, dosage, clinical outcome of interest, and methodology used to obtain the data (e.g., data obtained by PSG were analyzed separately from data obtained by sleep diary).

All meta-analyses were performed as per-treatment analyses using the random effects model. For most interventions, absolute effects of drug treatments are represented by the mean difference (MD) ± standard deviation (SD) of post-treatment drug versus post-treatment placebo. Meta-analyses for adverse events are presented as risk difference. The result of each meta-analysis is displayed as a forest plot. Pooled results are expressed as the total number of patients, MD and 95% confidence interval (CI) between the experimental treatment and placebo.

Interpretation of clinical significance for outcomes of interest was conducted by comparing the absolute effects of drug treatment to the clinical significance threshold previously determined by the task force for each outcome of interest. Interpretation of adverse events was based upon the risk difference and clinical expertise of the task force.

Strength of Recommendations

The GRADE approach (Grades of Recommendation, Assessment, Development and Evaluation) was used for the assessment of quality of evidence. For details on how the AASM uses GRADE to develop its clinical practice guidelines, refer to Morgenthaler et al.[53] The task force assessed the following three components to determine the direction and strength of a recommendation: quality of evidence, balance of beneficial and harmful effects, and patient values and preferences.

For the determination of the quality of evidence for an intervention, the task force used objective data whenever possible (e.g., PSG). When only subjective data were available (e.g., sleep diaries), this evidence was used to determine the overall quality of evidence. The decision to use objective data as the primary determinant of quality of evidence was based on the preference for an objective measure of physiologic changes for determining clinically significant efficacy, the standardization of sleep parameter measurements and reporting, and the current requirements of PSG data for FDA approval of hypnotic medications. The results of this assessment are presented as summary of findings tables for each intervention (see **Tables S1–S24** in the supplemental material).

The task force developed recommendation statements consistent with GRADE methodology based on the balance of the following factors:

1. Quality of evidence. Quality of evidence was based exclusively on the studies that could be included in meta-analyses. The task force determined their overall confidence that the estimated effect found in the literature was representative of the true treatment effect that patients would see, based on the following criteria: overall risk of bias (randomization, blinding, allocation concealment, selective reporting, and author disclosures); imprecision (when 95% CI cross the clinical significance thresholds); inconsistency (I^2 cutoff of 75%); indirectness (study population); and risk of publication bias (funding sources). The task force also considered the consistency of the supporting evidence (i.e. data the met inclusion criteria, but could not be included in the meta-analyses). However such evidence did not impact judgments regarding the quality of evidence or final recommendations.
2. Benefits versus harms. The task force determined if the beneficial outcomes of the intervention outweighed any harmful side effects based on the following criteria: meta-analysis (if applicable); analysis of any harms/side effects reported within the accepted literature; and the clinical expertise of the task force.
3. Patient values and preferences. The task force determined if patient values and preferences would be generally consistent, and if patients would use the intervention based on the body of evidence reviewed. These judgments were based on the clinical expertise of the task force members and any data published on the topic relevant to patient preferences.

Taking these major factors into consideration, and adhering to GRADE recommendations, the task force assigned a direction (for or against) and strength (STRONG or WEAK) for each recommendation statement.

Additional information is provided in the form of "Remarks" immediately following the recommendation statements, when deemed necessary by the task force. Remarks are based on the evidence evaluated during the systematic review, and are intended to provide context for the recommendations.

Approval and Interpretation of Recommendations

A draft of the guideline was made available for public comment for a two-week period on the AASM website. The task force took into consideration all the comments received and made revisions when appropriate. Based on recommendations from public comments, the task force decided to include data from clinicaltrials.gov, which allowed the development of a recommendation for the use of suvorexant. Due to a conflict of interest, Andrew Krystal did not participate in the development of the suvorexant recommendation. The final guideline was submitted to the AASM Board of Directors who approved these recommendations.

The recommendations in this guideline define principles of practice that should meet the needs of most adult patients, when pharmacologic treatment of chronic insomnia is indicated. This guideline should not, however, be considered inclusive of all proper methods of care or exclusive of other methods of care reasonably used to obtain the same results. A STRONG recommendation is one that clinicians should, under most circumstances, always be following when pharmacological treatment is indicated (i.e., something that might qualify as a quality measure). A WEAK recommendation reflects a lower degree of certainty in the appropriateness of the patient-care strategy and requires that the clinician use his/her clinical knowledge and experience, and refer to the individual patient's values and preferences to determine the best course of action. The ultimate judgment regarding propriety of any specific care must be made by the clinician, in light of the individual circumstances presented by the patient, available diagnostic tools, accessible treatment options and resources, as well as safety considerations.

Table 4—Summary of clinical practice recommendations and GRADE components of decision-making.

Treatment	Recommendation	Direction and Strength of Recommendation	Quality of Evidence	Benefits and Harms Assessment	Patients' Values and Preferences Assessment
Orexin receptor agonists					
Suvorexant This recommendation is based on trials of 10, 15/20, and 20 mg doses of suvorexant.	We suggest that clinicians use suvorexant as a treatment for sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Low	Benefits outweigh harms	The majority of patients would use this treatment (over no treatment), but many would not.
BZD receptor agonists					
Eszopiclone This recommendation is based on trials of 2 mg and 3 mg doses of eszopiclone.	We suggest that clinicians use eszopiclone as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Very low	Benefits outweigh harms	The majority of patients would use this treatment (over no treatment), but many would not.
Zaleplon This recommendation is based on trials of 10 mg doses of zaleplon.	We suggest that clinicians use zaleplon as a treatment for sleep onset insomnia (versus no treatment) in adults.	WEAK	Low	Benefits outweigh harms	The majority of patients would use this treatment (over no treatment), but many would not.
Zolpidem This recommendation is based on trials of 10 mg doses of zolpidem.	We suggest that clinicians use zolpidem as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Very low	Benefits outweigh harms	The majority of patients would use this treatment (over no treatment), but many would not.
Benzodiazepines					
Triazolam This recommendation is based on trials of 0.25 mg doses of triazolam.	We suggest that clinicians use triazolam as a treatment for sleep onset insomnia (versus no treatment) in adults.	WEAK	High	Benefits approx equal to harms	The majority of patients would use this treatment (over no treatment), but many would not.
Temazepam This recommendation is based on trials of 15 mg doses of temazepam.	We suggest that clinicians use temazepam as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Moderate	Benefits outweigh harms	The majority of patients would use this treatment (over no treatment), but many would not.
Melatonin agonists					
Ramelteon This recommendation is based on trials of 8 mg doses of ramelteon.	We suggest that clinicians use ramelteon as a treatment for sleep onset insomnia (versus no treatment) in adults.	WEAK	Very low	Benefits outweigh harms	The majority of patients would use this treatment (over no treatment), but many would not.
Heterocyclics					
Doxepin This recommendation is based on trials of 3 mg and 6 mg doses of doxepin.	We suggest that clinicians use doxepin as a treatment for sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Low	Benefits outweigh harms	The majority of patients would use this treatment (over no treatment), but many would not.
Trazodone This recommendation is based on trials of 50 mg doses of trazodone.	We suggest that clinicians not use trazodone as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Moderate	Harms outweigh benefits	The majority of patients would use this treatment (over no treatment), but many would not.
Anticonvulsants					
Tiagabine This recommendation is based on trials of 4 mg doses of tiagabine.	We suggest that clinicians not use tiagabine as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Very low	Harms outweigh benefits	The majority of patients would not use this treatment (over no treatment), but many would.
Over-the-counter preparations					
Diphenhydramine This recommendation is based on trials of 50 mg doses of diphenhydramine.	We suggest that clinicians not use diphenhydramine as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Low	Benefits approx equal to harms	The majority of patients would not use this treatment (over no treatment), but many would.
Melatonin This recommendation is based on trials of 2 mg doses of melatonin.	We suggest that clinicians not use melatonin as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Very low	Benefits approx equal to harms	The majority of patients would use this treatment (over no treatment), but many would not.
L-tryptophan This recommendation is based on trials of 250 mg doses of tryptophan.	We suggest that clinicians not use tryptophan as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults.	WEAK	High	Harms outweigh benefits	The majority of patients would use this treatment (over no treatment), but many would not.
Valerian This recommendation is based on trials of variable dosages of valerian and valerian-hops combination.	We suggest that clinicians not use valerian as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults.	WEAK	Low	Benefits approx equal to harms	The majority of patients would not use this treatment (over no treatment), but many would.

approx = approximately.

The AASM expects this guideline to have an impact on professional behavior, patient outcomes, and, possibly, health care costs. This clinical practice guideline reflects the state of knowledge at the time of publication and will be reviewed and updated as new information becomes available.

CLINICAL PRACTICE RECOMMENDATIONS

The following clinical practice recommendations are based on the systematic review and evaluation of evidence following the GRADE methodology. Remarks are intended to provide context for the recommendations. All meta-analyses and summary of findings tables are presented in the supplemental material. A summary of the recommendations and GRADE determinations is presented in **Table 4**. A summary of the recommendations, "critical" outcomes, and side effects is presented in **Table 5**.

It is essential that the recommendations which follow be interpreted within the appropriate context of clinical practice. Readers will note that all specific recommendations fall within the "weak" (for or against) classification of the GRADE system. **This should not be construed to mean that no sleep-promoting medications are clearly efficacious or indicated in the treatment of chronic insomnia**. Hypnotic medications, along with management of comorbidities and non-pharmacological interventions such as CBT, are an important therapeutic

Table 5—Summary of "critical" outcomes by indication.

Recommended for Treating Sleep Onset Insomnia	
Eszopiclone	**Sleep latency:** Mean reduction was 14 min greater, compared to placebo (95% CI: 3 to 24 min reduction); **Quality of sleep*:** Moderate-to-Large [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 2, "Harms" *This recommendation is based on trials of 2 mg and 3 mg doses of eszopiclone.*
Ramelteon	**Sleep latency:** Mean reduction was 9 min greater, compared to placebo (95% CI: 6 to 12 min reduction); **Quality of sleep*:** No improvement [b] in quality of sleep, compared to placebo; **Side effects:** See Recommendation 7, "Harms" *This recommendation is based on trials of 8 mg doses of ramelteon.*
Temazepam	**Sleep latency:** Mean reduction was 37 min greater, compared to placebo (95% CI: 21 to 53 min reduction); **Quality of sleep*:** Small [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 6, "Harms" *This recommendation is based on trials of 15 mg doses of temazepam.*
Triazolam	**Sleep latency*:** Mean reduction was 9 min greater, compared to placebo (95% CI: 4 to 22 min reduction); **Quality of sleep*:** Moderate [c] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 5, "Harms" *This recommendation is based on trials of 0.25 mg doses of triazolam.*
Zaleplon	**Sleep latency:** Mean reduction was 10 min greater, compared to placebo (95% CI: 0 to 19 min reduction); **Quality of sleep*:** No improvement [b] in quality of sleep, compared to placebo; **Side effects:** See Recommendation 3, "Harms" *This recommendation is based on trials of 5 mg and 10 mg doses of zaleplon.*
Zolpidem	**Sleep latency:** Mean reduction was 5–12 min greater, compared to placebo (95% CI: 0 to 19 min reduction); **Quality of sleep*:** Moderate [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 4, "Harms" *This recommendation is based on trials of 10 mg doses of zolpidem.*
Recommended for Treating Sleep Maintenance Insomnia	
Doxepin	**Total sleep time:** Mean improvement was 26–32 min longer, compared to placebo (95% CI: 18 to 40 min improvement); **Wake after sleep onset:** Mean reduction was 22–23 min greater, compared to placebo (95% CI: 14 to 30 min reduction); **Quality of sleep*:** Small-to-moderate [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 8, "Harms" *This recommendation is based on trials of 3 mg and 6 mg doses of doxepin.*
Eszopiclone	**Total sleep time:** Mean improvement was 28–57 min longer, compared to placebo (95% CI: 18 to 76 min improvement); **Wake after sleep onset:** Mean reduction was 10–14 min greater, compared to placebo (95% CI: 2 to 18 min reduction); **Quality of sleep*:** Moderate-to-Large [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 2, "Harms" *This recommendation is based on trials of 2 mg and 3 mg doses of eszopiclone.*
Temazepam	**Total sleep time:** Mean improvement was 99 min longer, compared to placebo (95% CI: 63 to 135 min improvement); **Wake after sleep onset:** Not reported; **Quality of sleep*:** Small [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 6, "Harms" *This recommendation is based on trials of 15 mg doses of temazepam.*
Suvorexant	**Total sleep time:** Mean improvement was 10 min longer, compared to placebo (95% CI: 2 to 19 min improvement); **Wake after sleep onset:** Mean reduction was 16–28 min greater, compared to placebo (95% CI: 7 to 43 min reduction); **Quality of sleep*:** Not reported; **Side effects:** See Recommendation 1, "Harms" *This recommendation is based on trials of 10, 15/20, and 20 mg doses of suvorexant.*
Zolpidem	**Total sleep time:** Mean improvement was 29 min. longer, compared to placebo (95% CI: 11 to 47 min. improvement); **Wake after sleep onset:** Mean reduction was 25 min greater, compared to placebo (95% CI: 18 to 33 min reduction); **Quality of sleep*:** Moderate [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 4, "Harms" *This recommendation is based on trials of 10 mg doses of zolpidem.*
Not Recommended for Treating either Sleep Onset or Sleep Maintenance Insomnia	
Diphenhydramine	**Sleep latency:** Mean reduction was 8 min greater, compared to placebo (95% CI: 2 min increase to 17 min reduction); **Total sleep time:** Mean improvement was 12 min longer, compared to placebo (95% CI: 13 min reduction to 38 min improvement); **Quality of sleep*:** No improvement [a] in quality of sleep, compared to placebo; **Side effects:** See Recommendation 11, "Harms" *This recommendation is based on trials of 50 mg doses of diphenhydramine.*
Melatonin	**Sleep latency:** Mean reduction was 9 min greater, compared to placebo (95% CI: 2 to 15 min reduction); **Quality of sleep*:** Small [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 12, "Harms" *This recommendation is based on trials of 2 mg doses of melatonin.*
Tiagabine	**Total sleep time:** Mean improvement was 1–7 min longer, compared to placebo (95% CI: 7 min reduction to 15 min improvement); **Wake after sleep onset:** Mean reduction was 1–9 min greater, compared to placebo (95% CI: 6 min increase to 25 min reduction); **Quality of sleep*:** No-to-Small [a] improvement in quality of sleep, compared to placebo; **Side effects:** See Recommendation 10, "Harms" *This recommendation is based on trials of 4 mg doses of tiagabine.*
Trazodone	**Sleep latency*:** Mean reduction was 10 min greater, compared to placebo (95% CI: 9 to 11 min reduction); **Wake after sleep onset:** Mean reduction was 8 min greater, compared to placebo (95% CI: 7 to 9 min reduction); **Quality of sleep*:** No improvement [d] in quality of sleep, compared to placebo; **Side effects:** See Recommendation 9, "Harms" *This recommendation is based on trials of 50 mg doses of trazodone.*
L-tryptophan	**Sleep latency:** Not reported; **Wake after sleep onset*:** Mean reduction was 10 min greater, compared to placebo (95% CI: 4 to 15 min reduction); **Quality of sleep*:** Small [e] improvement in quality of sleep, compared to placebo; **Side effects:** see Recommendation 13, "Harms" *This recommendation is based on trials of 250 mg doses of tryptophan.*
Valerian	**Sleep latency:** Mean reduction was 9 min greater, compared to placebo (95% CI: 0 to 18 min reduction); **Quality of sleep*:** Not reported; **Side effects:** See Recommendation 14, "Harms" *This recommendation is based on trials of variable dosages of valerian and valerian-hops combination.*

Drugs are listed alphabetically. All reported measures are based on polysomnographic data, unless otherwise noted. *Based on subjective reporting. [a] Based on Cohen's d: 0.2 = small effect; 0.5 = moderate effect; 0.8 = large effect. [b] Based on a 7-point Likert scale (1 = excellent, 7 = very poor). [c] Based on a 4-point scale (1 = good, 4 = poor). [d] Based on a 4-point scale (1 = excellent, 4 = poor). [e] Based on a 3-point scale (sleep quality index: 1 = low, 3 = high).

option for chronic insomnia. The strength of recommendations within the GRADE system are driven by the degree of confidence in a variety of factors related to the intervention including (1) the availability of specific data regarding efficacy; (2) the quality of that data, and (3) other considerations such as potential risks, impact of treatment, patient values and preferences, and perceived burden of treatment.

The existing data regarding sleep-promoting medications imposes limits on the degree of confidence as a result of several factors. These include: (1) a high degree of variability in the statistical information presented. Many studies, especially older studies, do not present results that meet the criteria for meta-analysis within GRADE and are, by necessity, excluded from formal analysis; (2) a significant degree of variability in sleep outcomes within and across studies. Such variability produces a "downgrading" of the quality of evidence within GRADE; (3) industry sponsorship. Very few clinical trials with adequate sample size have been sponsored by agencies outside of industry. As a result, the quality of evidence for a vast majority of available data is downgraded due to potential publication bias associated with such sponsorship; (4) a paucity of systematic data collection and analysis of treatment-emergent adverse events. Absent such information, it is difficult to assign a high degree of confidence to estimations of benefit:risk ratio; and (5) absence of outcome data (such as functional status or prevention of complications of chronic insomnia) that would inform judgments regarding the impact of therapy.

The strength (or weakness) of these recommendations speaks as much, or more, to the limitations of the data as it does to the relative benefits and risks of the treatments per se. Clinicians must continue to exercise appropriate judgement, based not only on the recommendations presented herein, but also on individual patient characteristics, comorbidities, and patient preferences in the prescribing of sleep-promoting medications and general management of chronic insomnia.

Finally, the literature review, meta-analyses, and recommendations are based only on FDA-approved doses. This should not be interpreted as a recommendation for the use of a specific dose in clinical practice. Numerous factors, including, but not limited to, age, sex, comorbidities, and concurrent use of other medications may affect dosage recommendations. Clinical judgment is necessary in determining appropriate dosage, on a patient-by-patient basis.

Orexin receptor agonists

Suvorexant for the Treatment of Chronic Insomnia

Recommendation 1: We suggest that clinicians use suvorexant as a treatment for sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 10, 15/20, and 20 mg doses of suvorexant.

Summary

Two RCTs[54,55] evaluated suvorexant for treatment of chronic primary insomnia. The statistical analyses and recommendation are based on data available on clinicaltrials.gov. The overall quality of evidence was low due to imprecision and risk of publication bias. The overall evidence for suvorexant was weakly in favor of its effectiveness for the treatment of sleep maintenance insomnia only. Objective reports of wake after sleep onset (PSG) showed clinically significant reduction at both 10 mg and 20 mg dosages. Subjective TST data demonstrated improvement, but failed to meet clinical significance. Objective reports (PSG) at the 10 mg and 15/20 mg dosages showed minimal improvements in sleep latency that failed to meet clinical significance. However objective reports (PSG) at 20 mg dose did show clinically significant reduction in sleep latency, suggesting that suvorexant may improve sleep onset at higher dosages. PSG sleep efficiency (SE) results demonstrate improvements that are near or above the level for clinical significance. PSG number of awakenings (NOA) was not statistically significantly reduced or increased in either study. Finally, sleep quality ratings showed minimal change.

Adverse events were assessed in both studies. Overall frequency of adverse events was not significantly increased versus placebo. There was no evidence of daytime residual or withdrawal symptoms. Therefore the task force judged the overall benefits to outweigh the potential harms. Based on their clinical judgement, the task force determined that the majority of patients would use suvorexant over no treatment.

See **Tables S1–S3** in the supplemental material.

Discussion

Two RCTs[54,55] evaluated suvorexant for treatment of chronic primary insomnia. However, data were not presented in a way that could be used for statistical analyses; therefore the statistical analyses and recommendation are based on data available on clinicaltrials.gov. Additional outcomes data from Herring 2012 and 2016 are discussed below as supporting evidence.

Herring 2012[55] evaluated adults 18–64 years of age with DSM-IV primary insomnia in a randomized placebo-controlled crossover study which included two 4-week trial periods. Sixty-two subjects received 10 mg suvorexant and 61 received 20 mg. Subjects underwent PSG at the end of week 4. Sleep diary data were also obtained. The primary endpoint was sleep efficiency; secondary endpoints included latency to persistent sleep and wake after sleep onset. Inclusion criteria were LPS > 20 min and WASO ≥ 60 min.

Herring 2016[54] conducted two randomized placebo-controlled parallel trials of 3 months each (i.e. trial 1 and trial 2). Only data from trial 1 were available for statistical analyses. Adults 18- to 64-years-old and adults > 65 with primary insomnia were included. Two-hundred fifty four and 239 patients were randomized to suvorexant 15/20 mg in the two trials, respectively. The dosages of interest for this analysis were 20 mg for younger adults and 15 mg for older adults. Data for the two dosages were pooled for analysis. Inclusion criteria were LPS > 20 min and WASO > 60 min. Sleep diary data was collected for all patients and a subset underwent PSG. Both studies reported data as difference between placebo and drug change from baseline.

SLEEP LATENCY: Herring 2012[55] found a reduction of 2.3 min (95% CI: 13.68 min lower to 9.08 min higher) for suvorexant

10 mg when compared to placebo (not considered clinically significant). The quality of evidence was low due to imprecision and potential publication bias. At the 20 mg dosage, a clinically significant reduction versus placebo of 22.3 min was reported (95% CI: 33.77 to 10.83 min lower). The quality of evidence was MODERATE due to potential publication bias. LPS in the first trial of Herring 2016[54] showed reductions of 8.1 min (95% CI: 13.85 to 2.35 min lower), and failed to meet clinical significance. The quality of evidence was low due to imprecision and potential publication bias. LPS in the second trial of Herring 2016 was not available for statistical analyses. However, published data show a reduction of 0.3 min, which also fails to meet the clinical significance threshold.

Subjective latency, reported as TSO in Herring 2016[54] trial 1, showed reductions at the pooled 15/20 mg dosages (−5.2 min; 95% CI: 0.3 to 10.1 min lower) that failed to meet the clinical significance threshold. The quality of evidence was moderate due to potential publication bias. Herring 2016[54] trial 2 reported reductions in TSO of 7.6 min, while TSO reported in the Herring 2012[55] study was reduced at both dosages (−3.0 min and −4.3 min at 10 mg and 20 mg, respectively); none of these changes met the clinical significance threshold.

Total sleep time: Herring 2016, trial 1, reported improvements in subjectively reported total sleep time of 10.6 min with 15/20 mg dosages (95% CI: 1.79 to 19.41 min higher), which did not meet the clinical significance threshold. The quality of evidence for this outcome was moderate based on potential publication bias.

PSG TST was reported only in the Herring 2012[55] investigation. At both 10 mg and 20 mg, clinically significant improvement was seen versus placebo (+22.3 min and +49.9 min, respectively).

Neither suvorexant 10 mg (+5.5 min) nor 20 mg (−1.8 min) produced statistically or clinically significant improvement in subjective TST versus placebo at 4 weeks (Herring 2012). In trial 2 of the 15/20 mg dosages (Herring 2016[54]), subjective TST was improved (+22.1 min), although the mean change falls below the clinical significance threshold.

Wake after sleep onset: Both studies reported PSG WASO. Herring 2012[55] found clinically significant reduction of WASO at both 10 mg and 20 mg (−21.4 min; 95% CI: 6.66 to 36.34 min lower; −28.1 min; 95% CI: 13.13 to 43.07 min lower, respectively). The quality of evidence was low due to imprecision and potential publication bias. Herring 2016,[54] trial 1, reported reductions of −16.6 min (95% CI: 8.33 to 24.87 min lower) with low quality of evidence due to imprecision and potential publication bias. Herring 2016[54] trial 2 reported a 31.1 min reduction in WASO. Reductions of subjective WASO in the two trials of 15/20 mg suvorexant in the Herring 2016[54] study did not meet clinical significance thresholds (−2.4 min and −7.7 min).

Quality of sleep: Sleep quality reductions were not statistically significant in either study.

Sleep efficiency: Herring 2012[55] found PSG SE improvement of +4.7% (95% CI: 0.97 to 8.43% higher) for 10 mg and +10.4% (95% CI: 13.13 to 43.07 min lower) for 20 mg, with low and moderate quality of evidence due to imprecision and potential publication bias. These values approximate (10 mg) or exceed (20 mg) the clinical significance threshold of 5%.

Number of awakenings: Number of awakenings showed no significant reduction in either study.

Overall quality of evidence: The overall quality of evidence for these studies was low due to imprecision and potential publication bias.

Harms: Neither study found a significant increase in one or more adverse events versus placebo for suvorexant in the 10–20 mg range. Rates of serious adverse events were negligible and not significantly different between suvorexant and placebo. Frequency of daytime somnolence was increased in the 15/20 mg doses (Herring 2012[55]: placebo 0.4%; 20 mg 4.9%. Herring 2016[54]: placebo = 3.4%; 15/20 mg = 5.1% [trial 1]; placebo = 3.1%; 15/20 mg = 8.4%). The degree of somnolence was reported to be typically mild to moderate and did not often result in discontinuation. Frequency of somnolence was noted to increase significantly in dose-dependent fashion at dosages exceeding FDA-recommended levels.

Assessments of withdrawal symptoms and daytime performance decrements did not reveal clinically significant findings in either domain. There was no evidence of the emergence of narcolepsy symptoms.

Patients' values and preferences: The task force determined that a majority of patients would likely use suvorexant compared to no treatment. This assessment reflects the task force's clinical judgment, based on suvorexant's efficacy for reduction of WASO and improvement in TST and SE and its relatively benign side effect profile.

BZD receptor agonists

Eszopiclone for the Treatment of Chronic Insomnia

Recommendation 2: We suggest that clinicians use eszopiclone as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 2 mg and 3 mg doses of eszopiclone.

Summary

Six RCTs evaluated eszopiclone 2 mg for the treatment of chronic primary insomnia.[56–61] The overall quality of evidence was downgraded to low due to imprecision and risk of publication bias. The evidence for eszopiclone 2 mg was weakly in favor of its efficacy for improving sleep onset disturbance and total sleep time. Meta-analysis data from three studies which reported objective sleep latency showed a clinically significant mean reduction in PSG sleep latency.[58,60,61] Four studies which evaluated subjective total sleep time demonstrated a significant mean increase versus placebo.[57–59,61] Assessment of PSG SE in

two studies[58,61] and subjective sleep quality in four studies[57,59–61] revealed improvements which fell just below the threshold for clinical significance. Measures of reduction in wake time after sleep onset and number of awakenings revealed trends toward improvement which fell below the defined level of clinical significance. Meta-analysis of adverse effects, derived from all six studies, indicated no significant differences versus placebo.

Six studies assessed the effects of eszopiclone 3 mg for treatment of chronic primary insomnia.[57,60–64] The quality of evidence for these studies as a whole was downgraded to very low due to significant heterogeneity, imprecision and potential publication bias. The collective evidence for eszopiclone 3 mg was weakly in favor of efficacy for improving sleep onset, total sleep time, sleep efficiency, number of awakenings and sleep quality. The meta-analysis data from three studies demonstrated clinically significant reduction in objective sleep latency.[60–62] Four studies likewise revealed clinically significant increase in mean subjective total sleep time.[57,61,63,64] PSG sleep efficiency, reported in two studies[61,62] also exceeded the threshold for clinically significant improvement, as did subjective sleep quality, which was reported in all six studies included in meta-analysis. A trend in the direction of reduced WASO was observed, but did not reach clinical significance. Insufficient data were available for meta-analysis of eszopiclone 3 mg adverse effects.

Overall, the benefits of eszopiclone 2 mg and 3 mg were judged to be greater than the potential harms. Based on clinical judgment, the task force determined that the majority of well-informed patients would use eszopiclone over no treatment. This judgement is based on the evidence of improvement in sleep latency, total sleep time, sleep efficiency and sleep quality, coupled with its low potential for adverse events.

See **Figures S1–S7**, **S68–S69**, and **Tables S4** and **S5** in the supplemental material.

Discussion

A total of nine studies were included in the meta-analyses for eszopiclone 2 mg and 3 mg.[56–64] Three of these studies included only older adults (> 65 years).[56,58,59] The remainder included younger adults, typically 21–65 years of age. Inclusion criteria for most of these studies required persistent subjective sleep latency > 30 min and TST < 6.5 h.[57–62] Ancoli-Israel and colleagues[56] studied 388 older adults for 12 consecutive weeks of nightly eszopiclone 2 mg. Inclusion criteria for this study specified TST < 6 h and WASO > 45 min. Outcome data were patient-reported. McCall and colleagues[58] also reported on two-week administration of 2 mg eszopiclone versus placebo to 254 to older adults. In addition to sleep latency and TST inclusion criteria, subjects were required to have WASO > 20 min. PSG was conducted on nights 1, 2, 13, and 14. Scharf and colleagues[59] administered 1 and 2 mg of eszopiclone or placebo nightly to 231 older adults for two weeks, employing nightly patient-reported data.

Erman and colleagues[57] evaluated multiple dosages of eszopiclone (1, 2, 2.5, and 3 mg versus placebo and an active control (zolpidem 10 mg) in 65 adult subjects (age 21–65) who received each intervention for two nights, followed by 3–7 day washout, in randomized sequences. PSG was conducted for the two nights on each treatment. The primary endpoint was latency to persistent sleep, with secondary endpoints of SE and WASO. Uchimura and colleagues[60] employed a similar crossover design with eszopiclone doses of 1, 2, and 3 mg, zolpidem 10 mg and placebo in 65 patients. PSG was conducted during each two-night intervention. Primary endpoints were objective latency to persistent sleep (LPS) and subjective SL. Zammit and colleagues[61] examined eszopiclone 2 and 3 mg vs. placebo for 44 consecutive nights, with PSG on nights 1, 15, 29. Patient-reported data were collected for nights 1, 15, 29, 43, and 44. Primary endpoint was PSG-defined LPS.

Krystal and colleagues[63] investigated six-month nightly use of eszopiclone 3 mg versus placebo in 788 adults. Patient-reported data were collected at weekly intervals. Similarly, Walsh and colleagues[64] reported on nightly use of eszopiclone 3 mg in 830 adults, with weekly patient-reported data. Finally, Boyle and colleagues,[62] in a study designed primarily to assess next-day driving skill, report subjective data from a single night of eszopiclone 3 mg versus placebo.

SLEEP LATENCY: Three studies assessed LPS as determined by PSG for eszopiclone 2 mg.[58,60,61] The McCall investigation[58] focused exclusively on older adults and demonstrated the greatest reduction in LPS. The mean reduction in LPS versus placebo for the three studies (−14.87 min; CI: −5.47 to −24.27 min) exceeded the threshold for clinical significance. The quality of evidence was LOW due to imprecision and potential publication bias.

All six trials of eszopiclone 2 mg reported subjective sleep latency.[56–61] As noted above, three of the six included only older adults. Mean difference from placebo fell slightly below the clinical significance threshold (−17.78 min; CI: −7.04 to −28.52 min). The quality of this evidence was low due to imprecision and potential publication bias.

Three studies investigated PSG LPS with eszopiclone 3 mg.[60–62] The mean difference in LPS (−13.63 min; CI: −3.7 to −23.56 min) fell below the clinical significance threshold. The quality of evidence was VERY LOW due to heterogeneity, imprecision and potential publication bias. Subjective SL with eszopiclone 3 mg was reported in four studies.[57,61,63,64] The mean difference exceeded the clinical significance threshold (−25.00 min; CI: −13.94 to −36.07 min). The greatest reductions were reported in the extended 6-month trials of Krystal and Walsh. Quality of evidence was low due to imprecision and potential publication bias.

Two additional studies not included in the meta-analysis reported subjective SL with eszopiclone 3 mg. Soares and colleagues[65] analyzed efficacy in perimenopausal/early menopausal women with sleep onset complaints. Joffe et al.[66] examined outcomes in perimenopausal/menopausal women who exhibited hot flashes and manifested either sleep onset or maintenance problems. The reductions in sleep latency versus placebo for these two studies (−15.7 and −17.8 min, respectively) were within the overall range found in the meta-analysis.

TOTAL SLEEP TIME: Only one eszopiclone study reported adequate objective total sleep time data. Therefore meta-analysis was not possible for this outcome at either dosage.[58] Four

studies included subjective TST for eszopiclone 2 mg.[57–59,61] The meta-analysis revealed a mean increase in TST of 27.53 min versus placebo, just below the threshold for clinical significance of 30 min. The quality of evidence was LOW due to imprecision and potential publication bias. The only study, noted above, which reported objective TST (in patients > 65 years) found an increase in TST of 28.6 min greater than placebo, consistent with the subjective results.

Four studies included adequate data for subjective TST meta-analysis for eszopiclone 3 mg.[57,61,63,64] These studies demonstrate substantially greater increases in TST at this dosage with a mean difference versus placebo of 57.1 min, exceeding the clinical significance threshold. The quality of evidence was moderate, due to potential publication bias.

The two studies of eszopiclone 3 mg in perimenopausal/early menopausal women revealed mean increases in subjective TST (versus placebo) of +66.5 min and +23.0 min.[65,66]

Wake after sleep onset: Two studies were included in the meta-analysis of objective WASO for eszopiclone 2 mg.[58,61] The mean reduction in WASO was 10.02 min greater than placebo, below the clinical significance level of 20 min for PSG data. The quality of evidence was rated as moderate due to potential publication bias. The confidence interval (−2.77 to −17.27 min) fell below the threshold.

Five studies reported adequate data for subjective WASO meta-analysis.[56–59,61] The mean difference versus placebo was below the threshold for clinical significance (−4.74 min; CI −11.87 to +2.39 min). The quality of evidence was moderate due to potential publication bias.

The data for PSG and patient-reported WASO with eszopiclone 3 mg demonstrated greater reduction of WASO than with 2 mg, but below clinical significance levels. The two studies including PSG WASO demonstrated a mean reduction of 14.69 min versus placebo (CI: −11.69 to −17.68 min).[61,62] Quality of evidence was moderate (potential publication bias). Subjective WASO for 3 mg was reported in four studies with mean reduction of 15.14 min (CI: −8.16 to −22.11 min). Quality of evidence was low due to imprecision and potential publication bias.

Krystal and colleagues[63] published an independent subgroup analysis of subjective WASO data from their 6-month nightly trial of 3 mg, in order to evaluate the impact of baseline WASO severity on outcome. They identified a positive relationship between baseline WASO severity and degree of improvement in WASO (as determined by eszopiclone/placebo difference) at all time points. The two investigations of menopausal women found eszopiclone-placebo mean differences for subjective WASO of 37.3 and 14.9 min, respectively.[65,66]

Quality of sleep: The meta-analysis for sleep quality with eszopiclone 2 mg included four studies and found a moderate effect size of +0.47 SMD (CI: +0.32 to +0.63 SMD).[57,59–61] The quality of evidence was moderate due to imprecision and potential publication bias. Sleep quality ratings for 3 mg, based on six studies, showed a large effect size of +0.82 SMD (CI: +0.41 to +1.24 SMD), although quality of evidence was very low due to imprecision, heterogeneity and potential publication bias.[57,60–64]

In addition to the studies included in meta-analysis, Soares and colleagues[65] reported statistically significant improvement in quality for eszopiclone 3 mg in their study of perimenopausal and postmenopausal women.

Sleep efficiency: Two studies reported PSG SE for eszopiclone 2 mg.[58,61] The mean improvement in SE of 4.83% fell below the significance threshold of 5%. (CI: 2.21 to +7.46%). For the 3 mg dosage, PSG SE exceeded the clinical significance threshold at 5.61%.[61,62] The quality of evidence for both doses was low due to imprecision and potential publication bias.

In studies outside the meta-analysis, Joffe[66] reported a 14.6% improvement versus placebo in SE with 3 mg.

Number of awakenings: The PSG NOA for 2 mg showed an increase of 0.12 awakenings based on two studies.[58,61] Evidence quality was MODERATE. Subjective NOA was based on four studies and likewise demonstrated no clinically significant difference from placebo. Evidence quality was moderate due to potential publication bias.[57–59,61]

Overall quality of evidence: The overall quality of evidence in the meta-analytic data from these studies was downgraded to very low for several reasons. Substantial heterogeneity across studies was noted for multiple outcomes. The data were also downgraded for imprecision, due to the relatively large confidence intervals, which cross the clinical significance thresholds for several outcomes. All of these studies were industry sponsored, resulting in further downgrading of evidence due to potential publication bias. The quality of evidence for individual outcomes ranged from moderate to very low. Therefore the overall quality of evidence was very low.

Harms: Sufficient data for meta-analysis of side effects was available only for the 2 mg eszopiclone dosage. Five side effects (dizziness, dry mouth, headache, somnolence and unpleasant taste) were included. Four studies examined dizziness with 2 mg eszopiclone and found no difference from placebo.[57,58,60,61] Two studies reported adequate data for dry mouth.[58,61] A +0.06 risk difference was reported for eszopiclone. For headache, four studies found essentially no difference between eszopiclone and placebo.[56,57,59,61] The same was true for next-day somnolence, based on five studies.[57–61] Finally, five studies found a +0.07 risk difference for unpleasant taste.[56–59,61]

Although meta-analysis was not possible for eszopiclone 3 mg, individual studies reported results which are consistent with those of the 2 mg dosage. Krystal and colleagues[63] reported numerically higher adverse event rates for somnolence (eszopiclone 9.1%; placebo 2.6%), unpleasant taste (26.1% versus 5.6%), dry mouth (6.6% versus 1.5%), and dizziness (9.8% versus 3.1%). Boyle[62] studied braking reaction time and other performance measures and found no difference between eszopiclone 3 mg and placebo. Walsh[64] reported significantly greater frequencies of adverse events including somnolence (eszopiclone: 8.8% versus placebo: 3.2%), unpleasant taste (19.7% versus 1.1%) and myalgia (6.0% versus 2.9%). No difference was seen on the Benzodiazepine Withdrawal Scale scores following discontinuation. Zammit[61] demonstrated no

impairment in digit symbol substitution at either 2 mg or 3 mg. Joffe and colleagues[66] reported a 15.2% risk for metallic taste, but placebo rate for this side effect was not identified. Soares and colleagues[65] found a significant increase in unpleasant taste with eszopiclone (17.6% versus 0.5%). Headache frequency was no different and report of dry mouth was slightly increased for eszopiclone (4.0% to 1.4%).

In summary, the task force found that there was weak evidence of efficacy in the treatment of sleep onset and maintenance insomnia, with limited or no consistent evidence of adverse events in excess of placebo, with the possible exception of unpleasant taste. Therefore, benefits were deemed to marginally outweigh harms.

PATIENTS' VALUES AND PREFERENCES: The task force determined that a majority of patients would likely use eszopiclone compared to no treatment. This assessment reflects the task force's clinical judgment, based on eszopiclone's efficacy for sleep onset and maintenance, and its relatively benign side effect profile.

Zaleplon for the Treatment of Chronic Insomnia

Recommendation 3: We suggest that clinicians use zaleplon as a treatment for sleep onset insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 10 mg doses of zaleplon.

Summary

Two RCTs meeting inclusion criteria investigated the use of zaleplon 5 or 10 mg in the treatment of chronic primary insomnia.[67,68] One of these reported only subjective outcomes[67], and one reported subjective and PSG outcomes.[68] No meta-analysis was possible for these studies, due to the manner of reporting results. The overall quality of evidence from these studies was downgraded to low due to imprecision and potential publication bias; both studies were industry supported. The overall evidence for zaleplon 10 mg support its efficacy for the treatment of sleep onset insomnia. At the 10 mg dose, one objective (PSG) study demonstrated a reduction in sleep latency from baseline that met the criterion for clinical significance, with an approximately 9.5 min difference from placebo. Subjective sleep latency, reported in one study, showed a non-significant change of −11.4 min. Subjective TST increased by approximately 21.5 min, but the difference from placebo was not statistically significant. WASO was not significantly different from placebo. Similarly, subjective sleep quality showed minimal differences from placebo. The overall evidence for zaleplon 5 mg did not support its efficacy for treatment of any insomnia symptoms, based on self-report studies only. No PSG studies at the 5 mg dose met inclusion criteria. Treatment-emergent adverse events showed no significant difference from placebo for zaleplon 10 mg or 5 mg, and only one study suggested a small increase in rebound using self-reported TST as the outcome.

Data from three additional studies of zaleplon 5–10 mg met our inclusion criteria but could not be included in meta-analyses because key outcome data were presented in insufficient detail.[69–71] However, the results of these three studies were consistent with those of the two studies presented above, in finding differences from placebo in subjective SOL but no significant differences in subjective TST or sleep quality.

Overall, the evidence for efficacy of zaleplon 10 mg is marginal, and the evidence for harm appears equivalent to placebo; therefore potential benefits minimally outweigh potential harms. The lack of evidence for efficacy of zaleplon 5 mg makes any potential benefits equivalent to its minimal potential harms.

Based on clinical judgment, the task force determined that the majority of well-informed patients would use zaleplon over no treatment. This judgement is based on the minimal evidence of improved sleep latency across PSG and self-report domains, coupled with a low potential for adverse events.

See **Tables S6** and **S7** in the supplemental material.

Discussion

Evidence from two RCTs which investigated the use of zaleplon 5 or 10 mg in the treatment of chronic primary insomnia was included in the main analysis of outcomes, although meta-analysis could not be performed because data were presented as medians, or as means with no standard deviation.[67,68,70] Subjects in each study met criteria for primary insomnia or insomnia associated with nonpsychotic mental disorder by either DSM-III-R or DSM-IV criteria, together with quantitative criteria for self-reported sleep disturbance (SOL ≥ 30 min, plus either subjective TST ≤ 6.5 h, WASO ≥ 30 min, or ≥ 3 awakenings) and associated daytime complaints. Walsh 2000[68] also required PSG LPS of > 20 min on two screening nights. Patients were 18–65 years of age[68,70] or 65 years and older.[67] Study designs included randomized, double-blind, placebo run-in with zaleplon 5–20 mg or placebo for 14–35 nights, followed by a 2–7 night placebo substitution. Walsh[68] used PSG outcomes, whereas the other two studies used self-report only. Data for zaleplon 20 mg were not considered here because this is not an FDA-approved dose.

SLEEP LATENCY: One study evaluated the impact of zaleplon 10 mg versus placebo on PSG sleep latency (SL).[68] This study showed a clinically significant 9.5 min reduction in mean sleep latency versus placebo (difference in median of 8.5 min) that approached the 10 min value considered clinically significant, and was judged by the task force to be sufficient evidence for making a recommendation. The CI (−0.19 to −18.80 min) crossed the clinical significance threshold, and therefore the quality of evidence was downgraded for imprecision. It was downgraded further due to the risk of publication bias since the study was industry-funded. The resultant quality of evidence is low.

Self-reported sleep latency was reported in one study,[68] which showed a reduction compared to placebo at the end of treatment (−11.40 min; CI: −26.36 to +4.56 min), which failed to meet the criterion for clinical significance. Hedner[67] also reported reductions in subjective sleep latency; however, the results could not be subject to meta-analysis, since the mean values were presented only in graphic form.

Additional studies not included in the primary analysis yield similar findings. Ancoli-Israel[69] conducted a randomized, double-blind, multi-center study of the efficacy of zaleplon 5 and 10 mg versus placebo in older adults with DSM-IV insomnia, using a similar study design to Hedner,[67] with self-report outcomes. This study reported significant differences between zaleplon 10 mg and placebo at both treatment weeks, and between zaleplon 5 mg and placebo at week 2 only. Elie 1999[70] reported significant differences on placebo at weeks 1–3 of treatment, with differences in the range of −8 to −15 min. Fry[71] reported a 28-day double-blind, placebo run-in and run-out study of adults with DSM-III-R insomnia. Median subjective sleep latency was signficantly different from placebo at weeks 1, 3, and 4 for zaleplon 10 mg, and at week 1 for zaleplon 5 mg. Because mean and standard deviation data were not reported, data from these two studies could not be formally evaluated in our meta-analysis.

TOTAL SLEEP TIME: The effects of zaleplon 10 mg on subjective TST were evaluated in one study.[68] Over the course of a five-week study, TST differed significantly from placebo only in week one, with a difference of 21.5 min between groups (CI: −5.6 to +48.6 min); this difference failed to meet the criterion for clinical significance. Quality of the evidence was downgraded to low due to imprecision and potential publication bias.

Objective TST was evaluated in 2 studies.[67,68] However, meta-analysis of these studies was not possible due to the manner of data reporting. These studies showed no consistent evidence of a zaleplon − placebo difference at the 10 mg or 5 mg dose of zaleplon. Mean/median differences in subjective TST at the end of treatment were in the range of +7 to +22.4 min in favor of zaleplon. The results of studies not included in our formal analysis[69–71] showed very similar findings for subjective TST, with inconsistent differences between placebo and zaleplon 10 mg.

The effects of zaleplon 5 mg versus placebo on subjective total sleep time were reported in one study.[67] No significant differences in median sleep time were found between zaleplon 5 mg and placebo across 2–4 weeks of treatment. The results of studies not included in our formal analysis[69–71] showed similar findings for subjective TST, with no differences between placebo and zaleplon 5 mg.

WAKE AFTER SLEEP ONSET: Objective WASO was evaluated in one study,[68] but failed to meet the criterion for clinical significance (−2.10 min; CI: −10.23 to +6.03 min). The quality of evidence was moderate, due to potential publication bias. Subjective WASO was not reported in any of the studies.

QUALITY OF SLEEP: Subjective sleep quality, evaluated on an ordinal 1–7 scale (1 = good, 7 = bad) was reported in one of the formally evaluated studies for both 5 mg and 10 mg.[67] At both dosages sleep quality improved (−0.10 points; CI: −0.27 to +0.07 points), but failed to meet the criterion for clinical significance. The quality of evidence for both doses was downgraded to moderate due to potential publication bias.

In three additional studies,[69–71] subjective sleep quality differed from placebo inconsistently at either dose; the majority of study weeks showed no difference between groups. Quality of evidence was downgraded for publication bias. Precision and heterogeneity could not be formally evaluated.

SLEEP EFFICIENCY: Neither PSG nor subjective sleep efficiency were formally evaluated in any of the studies reviewed here.

NUMBER OF AWAKENINGS: Number of awakenings were evaluated in the sole PSG study.[68] However, formal evaluation of findings was not possible. No data were presented in the paper, but NOA was reported not to differ between zaleplon 10 mg and placebo at any treatment week. Subjective NOA was evaluated in the two studies formally included in our evaluation but data were presented as median values and could not be included in meta-analyses. Hedner[67] reported a difference of uncertain clinical significance only at week 1 and Walsh[68] reported a difference only at week 3. Data from three additional studies not included in our formal analysis[69–71] showed no significant differences in NOA for either zaleplon 10 mg or zaleplon 5 mg at any study week.

OVERALL QUALITY OF EVIDENCE: As noted above, no meta-analyses could be conducted on data from studies of zaleplon. Some studies reported median data only, or mean values with no standard deviation, for some of the key outcomes. Still other studies presented data for key outcomes only in graphical form. The quality of evidence was downgraded for imprecision, due to the relatively large confidence intervals which cross the clinical significance thresholds for multiple outcomes. All of these studies were industry sponsored, resulting in further downgrading of evidence due to potential publication bias. The quality of evidence for individual outcomes ranged from moderate to low, therefore the overall quality of evidence was low.

HARMS: No meta-analysis was conducted on harms. Each of the individual studies showed no significant difference in the overall rate of treatment-emergent adverse events between zaleplon and placebo. Several symptoms related to the central nervous system were more frequent numerically among zaleplon treated patients, although these differences were not statistically significant due to the low overall incidence of adverse events. The most common adverse events in studies of zaleplon versus placebo included headache, asthenia, neurasthenia, pain, fatigue, and somnolence. There was no clear evidence of dose-dependent effects.

Several of the reviewed studies reported data from double-blind placebo runout periods. No significant withdrawal symptoms were noted on the Benzodiazepine Withdrawal Symptom Questionnaire.[70,71] The single PSG study noted no evidence of withdrawal upon discontinuation for the 10 mg dose. Evidence of discontinuation-related increases in subjective TST were noted at the zaleplon 5 and 10 mg dose in older adults, and for subjective SOL in older adults at the zaleplon 5 mg dose.[67,69] A small increase in NOA of the second discontinuation night was also noted with zaleplon 5 mg.[70] These differences were small in absolute magnitude and of doubtful clinical significance. Other studies did not report evidence of

rebound insomnia.[71] Categorically-defined rebound insomnia was not significantly different for zaleplon 5 mg or zaleplon 10 mg versus placebo.[69,70]

The task force found that there was weak objective evidence of efficacy for zaleplon 10 mg in the treatment of sleep onset insomnia that was just below criteria for clinical significance, and no consistent evidence for efficacy in TST. Likewise, there was no statistical evidence of adverse events in excess of placebo, although some treatment-emergent adverse events were numerically more prevalent in zaleplon groups. Evidence for withdrawal effects was weak, inconsistent, and unlikely to be clinically important. On balance, benefits were deemed to marginally outweigh harms.

PATIENTS' VALUES AND PREFERENCES: The task force determined that a majority of patients would likely use zaleplon compared to no treatment. This assessment reflects the task force's clinical judgment, based on zaleplon's efficacy for sleep onset, and its relatively benign side effect profile.

Zolpidem for the Treatment of Chronic Insomnia

Recommendation 4: We suggest that clinicians use zolpidem as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 10 mg doses of zolpidem.

N.B. Although 10 mg. was the recommended starting dosage for adults at the time of initial approval, the FDA has subsequently lowered the recommended starting dosage of immediate-release zolpidem products to 5 mg. Further, the FDA has recommended a reduction of starting dosage for extended-release forms of zolpidem from 12.5 mg to 6.25 mg.

Summary

Twelve RCTs evaluated zolpidem 10 mg for the treatment of chronic primary insomnia.[57,59,60,70,72–79] The overall quality of evidence was downgraded to very low due to significant heterogeneity, imprecision, and risk of publication bias. The evidence for zolpidem 10 mg was weakly in favor of its effectiveness for improving sleep onset, sleep maintenance, sleep quality, SE and TST. In addition, one paper evaluated the effectiveness of zolpidem extended release 6.25 mg[80] and one paper assessed zolpidem extended release 12.5 mg.[81]

Five studies examined the effects of zolidem 10 mg on objective sleep latency.[60,73,76,77,82] The mean reduction (vs. placebo) for PSG-determined latency to sleep exceeded the threshold for clinical significance. Ten studies presented patient-reported sleep latency data.[57,60,70,72–76,78,82] The mean reduction in subjective latency fell approximately at the significance threshold. Two studies[73,76] reported adequate objective TST data for meta-analysis and found that the mean improvement in TST also exceeded the clinical significance threshold. The same was true for subjective TST, based on eight studies.[57,70,73–76,78,82] Two studies[73,76] found that PSG-determined reduction in WASO was clinically significant. Six studies included adequate data for meta-analysis of subjective WASO[57,72,75,76,78,82]; the mean reduction fell below the clinical significance threshold. Six studies evaluating sleep quality reported moderately large improvement in this parameter based on SMD.[57,60,76,78,79,82] Improvement in PSG SE in the four studies included also exceeded the clinical significance threshold.[73,76,77,82] Number of awakenings (objective) fell below the clinical significance threshold.[77,82] Reduction in subjective number of awakenings also failed to meet the clinical significance threshold.

The single paper reporting on extended-release zolpidem 6.25 mg[80] found moderate reduction in PSG-determined WASO (based on only the first 6 h of sleep) and minimal improvement in LPS and SE at end-treatment in an elderly population. Overall quality of evidence from this report was LOW due to imprecision and potential publication bias. Data from the one study[81] on zolpidem extended-release 12.5 mg found moderate reduction in PSG LPS. Reduction in WASO was also moderate, while SE was not significantly different from placebo. Overall quality of evidence was LOW due to imprecision and potential publication bias.

Meta-analysis was conducted for amnesia, dizziness, headache, nausea, somnolence and "taste perversion" (altered or unpleasant taste) in studies employing zolpidem 10 mg. Small, but potentially significant increases in amnesia, dizziness and somnolence were reported with zolpidem.

Overall, the benefits of zolpidem 10 mg and extended-release zolpidem 12.5 mg were judged to be greater than the minimal potential harms. Benefits and harms were judged to be approximately equal for extended-release zolpidem 6.25 mg. It was determined by clinical judgement of the task force that the majority of well-informed patients would use zolpidem and extended-release over no treatment. This judgement is based on the evidence of improvement in sleep latency, total sleep time, WASO, sleep efficiency, and sleep quality, coupled with relatively low potential for adverse events. The data for efficacy of zolpidem extended-release 6.25 mg is minimal and inconclusive at best.

See **Figures S18–S27**, **S70–S75** and **Tables S8–S10** in the supplemental material.

Discussion

Twelve studies were included in the meta-analysis for zolpidem 10 mg.[57,60,70,72–79,82] Dorsey and colleagues[72] studied 141 menopausal or perimenopausal women who exhibited both insomnia (TST $\leq$ 6 h or WASO $\geq$ 1 h) and nocturnal hot flashes or sweats. Subjects received zolpidem 10 mg or placebo in a 4-week trial. Outcomes included patient-reported TST, SL, WASO, and NOA. Elie[70] investigated three dosages of zaleplon versus zolpidem 10 mg or placebo. The study included 615 adults with SL $\geq$ 30 min and either TST < 6.5 h or WASO $\geq$ 30 min or > 3 awakenings per night. Subjects received one of three zaleplon dosages, zolpidem 10 mg or placebo for 28 nights. Outcome data included subjective SL, QOS, TST and NOA. Erman[57] assessed 65 adults with reported sleep-onset insomnia and baseline PSG SL > 20 min and TST < 7 h or WASO > 20 min. Enrollees were administered eszopiclone at 4 dosages, zolpidem 10 mg and placebo in a randomized treatment sequence of 2 nights per intervention with intervening washout. Primary outcome was PSG-determined LPS with secondary measures including SE, WASO and NOA. Hermann[73] administered

zolpidem 10 mg or placebo for two weeks to 21 adults with difficulty initiating or maintaining sleep. PSG was conducted on the final treatment night with reported outcomes including SL, TST, SE and WASO.

Perlis[75] evaluated 199 subjects with primary insomnia (SL ≥ 45 min or TST ≤ 6 h) with zolpidem 10 mg or placebo. Subjects were instructed to take the medication 3–5 times per week as needed over a twelve-week period. Sleep diary outcomes included SL, TST, WASO and NOA. Jacobs[74] compared zolpidem 10 mg, cognitive behavior therapy and placebo in 63 adults with primary sleep-onset insomnia (SL > 1 h on ≥ 3 nights/week). Subjects received zolpidem for 28 days, followed by taper. Primary outcome was patient-reported sleep latency with secondary outcomes of SE and TST. Randall[76] investigated the efficacy of zolpidem 10 mg (5 mg for subjects 65–70 years) over an eight-month period in 91 subjects (age 23–70 years) with screening PSG SE < 85%. Patient-reported outcomes and PSG data at one and eight months included SL, TST, WASO and SE. Scharf[82] evaluated 75 adults for five weeks with zolpidem 10 mg, 15 mg or placebo. Inclusion criteria included SL ≥ 30 min or TST < 6 h. Subjects underwent sleep studies on the first two nights of each treatment week. Primary outcomes were defined as LPS and SE.

Staner[79] assessed the effects of three drugs, including zolpidem 10 mg, in a driving simulation study of 23 adults with recurrent SL > 30 min or WASO > 60 min. Sleep quality data was reported. Uchimura[60] compared zolpidem, eszopiclone and placebo in a crossover design as described in the eszopiclone section. Walsh[78] compared the efficacy of zolpidem 10 mg to trazodone 50 mg and placebo in 278 adults with insomnia characterized by frequent SL > 30 min and TST 4–6 h. Subjective sleep latency and TST were reported. Ware[77] assessed rebound insomnia in zolpidem 10 mg, triazolam 0.5 mg and placebo. Ninety-nine subjects with baseline PSG-determined LPS > 20 min and TST 4–7 h took zolpidem 10 mg, triazolam or placebo for 28 consecutive days. PSG LPS, SE, TST, and WASO were evaluated.

Two studies reported on extended-release (ER) zolpidem. Roth[81] assessed zolpidem ER 12.5 mg in 212 adults with insomnia who reported > 1 h WASO at least 3 nights per week. Patients received zolpidem or placebo nightly for 3 weeks in a parallel group design. Walsh[80] studied 205 elderly adults with insomnia with the same inclusion criteria and design, employing a 6.25 mg dose of zolpidem ER versus placebo.

Fourteen additional studies met inclusion criteria but could not be included in meta-analysis due to inadequate data sets.[71,83–95] Pertinent results from these studies are noted independently of meta-analysis results.

Sleep latency: Five studies included adequate data for PSG SL meta-analysis.[60,73,76,77,82] The mean difference from placebo of −11.65 min exceeded the clinical significance threshold. The 95% CI of −4.15 to −19.15 min crossed the clinical significance threshold and was therefore considered imprecise. Heterogeneity was high. With potential publication bias as well, the quality of evidence was rated as very low.

Ten of the twelve studies used in meta-analysis reported subjective SL.[57,60,70,72–76,78,82] The improvement in sleep latency versus placebo was at the significance threshold (mean difference: 19.55 min; CI: −14.2 to −24.9 min). Evidence quality was very low due to imprecision, heterogeneity and potential publication bias.

Six additional studies assessed sleep latency outcomes with zolpidem.[88–90,92,94,95] These studies varied significantly with regard to drug preparation, dosage, mode of administration and methodology, rendering comparisons between them or to the meta-analytic data difficult. Four of the six evaluated sublingual zolpidem, primarily for treatment of middle-of-the-night (MOTN) awakenings. Roth and colleagues[88] reported results of a three-way crossover study of zolpidem sublingual 1.75 mg, 3.5 mg and placebo. Zolpidem reduced both objective (latency to persistent sleep) and subjective latency to sleep (SL) following MOTN awakenings (PSG: 1.75 mg: −11.2 min versus placebo; 3.5 mg: −18.4 min/subjective: 1.75 mg: −11.83 min versus placebo; 3.5 mg: −15.23 min). Roth[89] also reported reduced subjective latencies following MOTN awakenings with sublingual zolpidem 3.5 mg over a 28-day trial. Zammit[95] administered immediate release oral zolpidem 10 mg, zaleplon 10 mg or placebo to subjects with sleep maintenance insomnia following induced MOTN awakenings Zolpidem reduced PSG latency to persistent sleep following the awakenings (−30.5 min versus placebo). Staner[90] compared the effects of sublingual zolpidem 10 mg to immediate release oral zolpidem on PSG initial sleep latency and reported shorter latency to persistent sleep with the sublingual preparation (−10.28 min) versus the oral preparation. Walsh[94] investigated subjective SL in an 8-week trial of as-needed zolpidem 10 mg (3–5 times per week). For medication nights only, end treatment SL for the zolpidem 10 mg group was 12.6 min less than the placebo group.

Walsh[80] investigated the effects zolpidem ER 6.25 mg and found reduction of PSG LPS of 13.0 min. Roth[81] reported a decrease in PSG LPS of 8.2 min versus placebo at end of treatment with zolpidem ER 12.5 mg.

Total sleep time: Two studies[73,76] were included in the meta-analysis of PSG-determined TST. Mean reduction in TST met the clinical significance threshold at +28.91 min, however the 95% CI crossed the threshold (CI: +10.85 to +46.97 min). The quality of evidence was downgraded to LOW due to imprecision and potential publication bias. Eight studies reported adequate data for meta-analysis of patient-reported TST.[57,70,73–76,78,82] The mean difference for subjective TST from these studies exceeded the significance threshold (+30.04 min; CI: +15.12 to +44.96 min). Quality of evidence was low due to imprecision and potential publication bias.

Six additional studies presented TST data which was not sufficient to be included in the analysis.[71,83,84,88,92,95] Allain and colleagues[83] evaluated zolpidem 10 mg administered on an as-needed basis over a four week period. When only drug nights were included in analysis, zolpidem produced a statistically significantly greater increase in subjective TST versus placebo (+19.9 min). Cluydts[84] and Hajak[85] found no difference in subjective TST with nightly versus intermittent (5/7 nights) use of zolpidem 10 mg, both of which produced numerical improvement (+11.3 and +16.9 min, respectively). In a study designed primarily to address potential rebound insomnia following

four weeks of treatment with zaleplon, zolpidem or placebo, Fry[71] reported substantially greater improvement in patient-reported TST with zolpidem 10 mg versus placebo (+28.2 min). In another study of rebound insomnia, Voshaar[92] compared zolpidem 10 mg to temazepam 20 mg (without placebo control) administered nightly for four weeks. The two drugs produced improved TST without significant difference between the two. Finally, in two studies Roth[89] and Zammit[95] investigated effects of zolpidem following MOTN awakenings. Roth and colleagues compared sublingual zolpidem 1.75 mg and 3.5 mg to placebo. Both dosages produced greater TST following awakening as compared to placebo (+14.7 min and +25.9 min, respectively). Zammit administered zolpidem 10 mg following MOTN awakening and reported TST after awakening 30 min greater than placebo.

TST data were not reported in the extended-release studies.

Wake after sleep onset: Two studies reported adequate data for meta-analysis of PSG WASO.[73,76] These studies yielded a mean difference from placebo of −25.46 min (CI: −17.94 to −32.99 min). This exceeds the threshold for clinical significance. The quality of evidence was LOW due to imprecision and potential publication bias.

Zolpidem ER 12.5 mg reduced WASO by 20 min greater than placebo at treatment conclusion, although this was based on only the first 6 h of sleep.[81] Comparison of changes from baseline in this study, however, suggested smaller differences between drug and placebo. Walsh,[80] using the same selective sample of 6 h, found WASO 13.0 min less in the zolpidem ER 6.25 mg group than in the placebo group. Given the sampling of only 6 h, it is difficult to clearly determine whether or not these agents would fulfill the criterion for clinical significance, which is based on an entire night of sleep.

Six studies assessed patient-reported WASO.[57,72,75,76,78,82] The mean difference fell below the level of clinical significance at −13.57 min (CI: −7.30 to −19.84 min). Quality of evidence was low due to heterogeneity and potential publication bias.

Quality of sleep: Six studies included sleep quality data.[57,60,76,78,79,82] The meta-analysis produced a standardized mean difference of +0.64 (CI: +0.03 to +1.26 SMD), suggesting moderate overall improvement in patient-reported sleep quality. Quality of evidence was very low due to imprecision, heterogeneity and potential publication bias.

Sleep efficiency: PSG sleep efficiency was reported in four studies.[73,76,77,82] The mean difference favored zolpidem (+6.12%; CI: +4.39 to +7.85%), but did not exceed the clinical significance threshold. Quality of evidence was low.

In the Roth[81] study of zolpidem ER 12.5 mg, PSG SE was 3.9% better with zolpidem than placebo. Walsh[80] found a difference of 2.4% on favor of zolpidem ER 6.25 mg. Neither value exceeds the clinical significance threshold.

Number of awakenings: PSG-determined number of awakenings was reported by Scharf[82] and Ware.[77] The mean difference from placebo was −0.95 awakenings (CI: −0.49 to −1.41 awakenings), which fails to meet the clinical significance threshold. Quality of evidence was moderate. Subjective awakening was reported in six studies.[70,72,73,75,78,82] Mean reduction versus placebo was −0.31 awakenings (CI: −0.17 to −0.45 awakenings), which also fails to achieve clinical significance. Evidence quality was low due heterogeneity and potential publication bias.

Overall quality of evidence: The overall quality of evidence in the meta-analytic data from these studies was downgraded to very low for several reasons. Substantial heterogeneity across studies was noted for multiple outcomes. The data were also downgraded for imprecision, due to the relatively large confidence intervals which cross the clinical significance thresholds for several outcomes. All of these studies were industry sponsored, resulting in further downgrading of evidence due to potential publication bias. The quality of evidence for individual outcomes ranged from moderate to very low, therefore the overall quality of evidence was very low.

Harms: Meta-analysis for adverse effects of zolpidem was possible for six side effects: amnesia, dizziness, sedation, headache, nausea, and taste perversion (altered or unpleasant taste). Two studies[70,82] included data on amnesia and found a minimal difference from placebo (0.03 risk difference). A small increase in risk (0.06 risk difference) was identified for dizziness, based on analysis of four investigations.[57,60,72,82] Risk for headache was mildly increased in the zolpidem group (0.07 risk difference).[57,72,78] Minimal difference was observed in the risk for nausea (0.02 risk difference),[57,82] and somnolence had a slightly higher risk (0.04), based on six studies.[57,60,70,72,78,82] Risk for taste perversion was low and approximately equal in both groups.[60,70]

Numerous studies have evaluated rebound insomnia after discontinuation of zolpidem.[68,70,71,73,75,77,80,82,86,92] Some of these studies found no evidence of rebound after varying durations of nightly or intermittent use, for up to six months.[68,73] Other investigations reported evidence of rebound, limited primarily to night 1 following discontinuation.[70,71,80,81]

Evaluation of daytime improvement and impairment was limited. Dorsey[72] reported improvement in sleep-related difficulty with daytime function. Hajak[85] described marked improvement in quality of life ratings with both nightly and intermittent use. Morning alertness and performance impairment were tested in several studies. Roth[81] and Walsh[80] found no evidence of impairment on digit symbol substitution test (DSST) or Rey auditory-verbal learning test (RAVLT) after zolpidem modified-release 12.5 mg. Scharf[82] reported no impairment on DSST or digit symbol copying. Staner[79] found no indication of impairment in a driving simulation study after seven consecutive nights of zolpidem 10 mg. Zammit[95] formally assessed sleepiness following administration of zolpidem 10 mg following MOTN awakening. Subjects showed significantly reduced PSG latencies versus placebo at 4, 5, and 7 h following administration. This was accompanied by impairment on DSST at 4 and 5 h.

In summary, the task force found that there was weak evidence of efficacy in the treatment of sleep onset and maintenance insomnia, with limited evidence of mild adverse events

in excess of placebo, with the possible exception of excessive sleepiness following administration of higher dosages (10 mg) less than 8 h prior to awakening. Therefore, benefits were deemed to marginally outweigh harms.

PATIENTS' VALUES AND PREFERENCES: The task force determined that a majority of patients would likely use zolpidem compared to no treatment. This assessment reflects the task force's clinical judgment, based on zolpidem's efficacy for sleep onset and maintenance, and its relatively benign side effect profile.

Benzodiazepines

Triazolam for the Treatment of Chronic Insomnia

Recommendation 5: We suggest that clinicians use triazolam as a treatment for sleep onset insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 0.25 mg doses of triazolam.

Summary

Because only one study[96] contained data of sufficient quality, meta-analysis was not performed. The quality of evidence for this study was high. This study, consisting of patient-reported data, showed a modest decrease in subjective SL, which fell below the clinical significance threshold. Two additional studies, which did not contain data suitable for meta-analysis, reported statistically significant reductions in subjective SL with triazolam 0.25 mg versus placebo.[97,98] Roehrs[96] found an increase in TST, although the mean change fell below the range of clinical significance. WASO was not reported, while sleep quality showed mild to moderate *reduction* versus placebo. Number of awakenings was insignificantly decreased.

No meta-analysis of harms was possible. Given the marginal evidence for efficacy in improving sleep onset, coupled with limited evidence regarding harms, the task force judged the harms to be approximately equal to the benefits. Based on its clinical judgement, the task force determined that, in light of the evidence for efficacy for sleep onset and the absence of information regarding harms, the majority of patients would be likely to use triazolam compared to no treatment.

See **Table S11** in the supplemental material.

Discussion

Roehrs[96] studied 32 adults with insomnia in a complex design which began with 11 days in which subjects received either triazolam or placebo nightly, "as needed" or every third night. This was followed by 14 nights in which subjects chose to self-administer treatment, with placebo (week 1) or triazolam 0.25 (week 2).

Thirteen additional studies met general inclusion and exclusion criteria.[97–109] These studies were highly varied in design, many utilizing interval scales (as opposed to specific numeric values) for reporting of sleep outcome variables. Some did not include a placebo comparison. Many included dosages which are higher than current recommended dosages. Therefore, only those studies which contained pertinent data are discussed.

Bowen[100] compared triazolam 0.5 mg, flurazepam 30 mg and placebo in 120 insomnia outpatients. The two-night crossover comparison of triazolam 0.5 mg and placebo included only 18 subjects, who completed morning sleep questionnaires. Elie[97] evaluated triazolam 0.125 mg (with upward dosage adjustment to 0.25 mg during the study period, as indicated) versus zopiclone and placebo in 48 elderly (60–90 years) subjects. Subjects received one of three interventions nightly for three weeks in a parallel group design. Outcome variables were patient-reported. Greenblatt[103] reported an RCT of 6 nights baseline placebo administration followed by triazolam 0.5 mg for 7–10 nights in a total of 60 subjects with sleep onset or maintenance insomnia. Outcome data were derived from subjective reports. Hajak[104] treated 1,507 subjects with sleep onset or maintenance insomnia with triazolam 0.25 mg, zopiclone or placebo. The triazolam versus placebo comparison groups totaled 605 subjects, who received drug or placebo for 28 consecutive nights and reported sleep variables on visual analog scales.

Monti[106] assessed 24 chronic insomnia subjects with triazolam 0.5 mg, zolpidem and placebo in a 27-night trial, with PSG on nights 4/5 and 15/16 and 29/30. Reeves[98] evaluated 37 geriatric subjects (> 60 years) with triazolam 0.25 mg, flurazepam or placebo in a 28 day trial. The triazolam and placebo groups included 28 subjects who completed daily sleep diaries. Rickels[107] studied 50 subjects with sleep onset or maintenance insomnia who received either triazolam 0.5 mg or placebo for 7 days. Outcome data were subjective ratings and interval scales. Scharf[108] administered triazolam 0.5 mg, quazepam or placebo to 65 chronic insomnia subjects. After placebo run-in, participants received nightly drug or placebo for 9 nights, followed by 14 nights of every-other-night administration. Outcomes were patient reported rating scales.

SLEEP LATENCY: In the only study with adequate data for meta-analysis, Roehrs[96] found a small reduction in subjective SL (−9.2 min; CI: −22.3 to +3.9 min) which fell below clinical significance. Quality of evidence for these data was high.

Monti[106] found no significant differences between triazolam 0.5 mg and placebo for PSG SL at any time point.

Elie[97] found larger reductions in subjective ratings of SL for triazolam 0.125–0.25 mg versus placebo. Hajak[104] found no significant difference from placebo in SL "response rate" (SL reduction of ≥ 15 min) for triazolam 0.25 mg. In contrast, Reeves[98] found triazolam 0.25 mg statistically superior to placebo for SL in a geriatric population on subjective ratings of "how much [the drug] helped." Bowen[100] found triazolam 0.5 mg to be statistically significantly better than placebo on interval ratings for reduction of sleep onset time. Greenblatt[103] reported sleep diary reductions from baseline placebo levels of 55 min and 24 min in two separate triazolam 0.5 mg groups. Rickels[107] reported similar subjective improvement on ratings of sleep induction for triazolam 0.5 mg.

TOTAL SLEEP TIME: Roehrs[96] observed a moderate increase in subjective TST (+25.20 min; CI: −9.12 to +59.52 min). This fell below the clinical significance threshold of 30 min and was not statistically different from placebo. Quality of evidence was moderate due to imprecision.

In additional studies, Hajak[104] found no significant difference between triazolam 0.25 mg and placebo in "percentage of responders" for TST (defined as > 20% increase in TST). Ratings for improvement in TST were significantly better for triazolam 0.25 mg than placebo in the Reeves[98] geriatric study.

Monti[106] observed a statistically significant increase in objective TST with triazolam 0.5 mg (+16 min versus placebo). Bowen[100] found that triazolam 0.5 mg was significantly preferred to placebo for sleep duration. In two separate triazolam 0.5 mg groups, Greenblatt[103] noted increases in subjective TST of 1.02 and 0.76 h. Rickels[107] reported that triazolam 0.5 mg was rated as significantly superior to placebo for sleep duration. Scharf[108] noted significantly greater subjective improvement in interval ratings of TST with both daily and every other night administration of triazolam 0.5 mg.

WAKE AFTER SLEEP ONSET: No studies reported data on WASO.

QUALITY OF SLEEP: Roehrs,[96] using a 4-point scale (1 = good, 4 = poor), found small improvements in QOS (−0.37 points; CI: −0.66 to −0.07 points), which was not considered to be clinically significant. Quality of evidence was high.

Elie[97] found no significant difference between triazolam (0.125/0.25 mg) and placebo in an elderly population with respect to QOS. Likewise, Hajak[104] reported no statistically significant difference between triazolam 0.25 and placebo. Reeves[98] demonstrated significant improvement in QOS for triazolam 0.25 mg versus placebo in a geriatric population.

SLEEP EFFICIENCY: Sleep efficiency was not reported by any study.

NUMBER OF AWAKENINGS: Roehrs[96] reported a reduction in NOA of 0.37 (CI: −1.7 to +0.96 awakenings), which did not meet the clinical significance threshold. Quality of evidence was LOW due to significant imprecision.

Hajak[104] noted no significant difference from placebo in the percentage of "responders" (reduction of NOA to ≤ 3) with triazolam 0.25 mg. However, Reeves[98] did find a statistically significant reduction at this dosage in ratings for NOAs. Bowen[100] observed a statistically significant reduction in subjective ratings for NOA with triazolam 0.5 mg versus placebo. Greenblatt[103] also reported reductions of 0.58 and 0.89 patient-reported awakenings from placebo baseline in two groups administered 0.5 mg triazolam.

OVERALL QUALITY OF EVIDENCE: The overall quality of evidence for the triazolam data, based on the single study meeting criteria for meta-analysis, was HIGH.

HARMS: Insufficient data were available for meta-analysis of adverse events associated with triazolam 0.25 mg. Very little systematic analysis of adverse effects is available. Hajak[104] reported that "speech disorder" was the only adverse effect, among many, to be significantly more frequent in the triazolam group than in the placebo condition.

PATIENTS' VALUES AND PREFERENCES: The task force determined that a majority of patients would be likely to use triazolam compared to no treatment. This assessment reflects the task force's clinical judgment, based on weak evidence for triazolam's efficacy and the absence of information regarding harms.

Temazepam for the Treatment of Chronic Insomnia

Recommendation 6: We suggest that clinicians use temazepam as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 15 mg doses of temazepam.

Summary

Three RCTs investigated the use of temazepam in the treatment of chronic primary insomnia.[110–112] These studies provide a limited assessment of temazepam in that they included small sample sizes of 19, 20, and 34 subjects, respectively. The overall quality of evidence from these studies is moderate. Meta-analyses for temazepam 15 mg were conducted for SL, TST and sleep quality. Two studies[110,112] were included in the meta-analysis of SL. These showed a reduction in subjective SL which exceeded the threshold for clinical significance. Meta-analysis of TST showed improvement in subjective TST which exceed the threshold for clinical significance. There were insufficient data for meta-analysis of WASO. One study of objective WASO revealed a clinically significant reduction. Subjective and objective SE was significantly increased, based on limited data from secondary studies. There was evidence for marginal improvement in sleep quality of 0.25 standard deviations. This was not a clinically significant difference from placebo and falls below the threshold for clinical significance. There were minimal data on adverse effects, and the available data do not suggest a high frequency of treatment-emergent adverse events (TEAEs).

Meta-analysis for temazepam 30 mg was not possible for any sleep outcomes. Data from individual studies are reported below.

In summary, meta-analysis data are available for temazepam 15 mg only. These data, coupled with data from secondary studies not adequate for meta-analysis, demonstrate efficacy for temazepam 15 mg in improving subjective and possibly objective sleep latency, subjective and objective TST, and objective WASO (the latter based on a single study). Temazepam 30 mg appears to have significant efficacy for improving subjective sleep latency and TST. The data also support a clinically significant effect for both 15 mg and 30 mg on subjective NOA, although data for objective NOA at 20 mg revealed no significant effect.

Given the significant improvements in patient-reported SL and TST, coupled with additional data derived from secondary studies (see below), the task force judged that the benefits of temazepam 15 mg appear to be greater than the potential harms. Based on its clinical judgement, the task force determined that, in light of the evidence for efficacy and minimal evidence for

harms, the majority of well-informed patients would be likely to use temazepam compared to no treatment.

See **Figures S28–S30** and **Tables S12** and **S13** in the supplemental material.

Discussion

Evidence from three RCTs which investigated the use of temazepam in the treatment of insomnia was included in the statistical analysis.[110–112]

Glass[110] evaluated 19 subjects 70 years of age or older who met DSM-IV diagnostic criteria for primary insomnia. Subjects underwent a crossover study of two weeks of treatment with placebo, temazepam 15 mg, or diphenhydramine 50 mg with randomized order of administration. Sleep was assessed using diary-derived variables. Adverse effects were recorded and daytime impairment was systematically assessed using the digit symbol substitution task (DSST), the manual tracking task (MTT), and the morning-after memory impairment, using a free-recall procedure.

Wu[112] assessed 71 patients with DSM-IV diagnosed insomnia who were randomized to one of four groups (CBT-I alone, CBT-I plus pharmacotherapy with temazepam 15 mg, pharmacotherapy alone, or placebo). For the purpose of this analysis, pharmacotherapy alone was compared to placebo (n = 34). Subjects received 8 weeks of treatment. End-of-treatment PSG and patient diary data for SL, TST and SE were compared.

Hindmarch[111] studied 20 individuals with "a history of nighttime medication for insomnia." No additional diagnostic information was provided. Subjects were randomized to receive temazepam 15 and 30 mg or placebo for a single night using a within-subjects crossover design. Outcome was assessed using the Leeds Sleep Evaluation Questionnaire which consisted of Visual Analogue Scale (VAS) ratings of "Ease of Falling Asleep" and "Quality of Sleep." Adverse effects were not reported, but daytime sedation was assessed with a Choice Reaction Time task, the Critical Flicker Fusion Test, and "Ease of Awakening" and "Integrity of Behavior Following Wakefulness" items from the Leeds scale.

Six additional studies which included temazepam-placebo comparisons were reviewed.[92,113–117] Cuanang[113] studied 60 adult "outpatients with insomnia." Parallel group design included three groups: temazepam 20 mg, temazepam 10 mg and placebo. Subjects received treatment or placebo for five nights. Patient-reported data including sleep quality ("better, same or worse"), SL, and TST were collected each morning. Fillingim[114] evaluated 75 adult patients with difficulty initiating (SL > 30 min) and maintaining (≥ 1 awakening with difficulty returning to sleep) sleep and TST < 6 h. Subjects received temazepam 30 mg, glutethimide or placebo in parallel group design for four nights. Outcomes included patient-reported estimates of SL, TST, NOA and QOS. Heffron[115] reported on 55 "insomnia outpatients" who received temazepam 30 mg or placebo in parallel groups for four nights. Subjects reported SL, TST, NOA and QOS. Tuk[116] studied 21 "primary sleep-onset insomnia" patients in a within-patient crossover study. Subjects received a single night of placebo and a single night of temazepam 20 mg with one-week intervening washout. PSG was conducted on each of the two nights. SL, TST, WASO and SE were reported. Voshaar[92] assessed 85 individuals with DSM-III-R primary insomnia in a within-subjects crossover design including temazepam 20 mg, zolpidem and placebo. A single-blind placebo period of four days was followed by 28 days of active treatment with zolpidem or temazepam. Data are presented as means for the placebo period and active treatment period for each sleep outcome. Wilson[117] conducted an actigraphic evaluation of 38 subjects with "complaints of poor sleep." Subjects were randomized to one of two crossover designs, each of which included two weeks of placebo and two weeks of temazepam 20 mg. Subjective results from patient diaries as well as actigraphic results were averaged over the respective periods.

SLEEP LATENCY: The meta-analysis for subjective SL, based on two studies[110,112] of temazepam 15 mg revealed a mean reduction of 20.06 min (CI: −1.07 to −39.05 min). Quality of evidence was moderate due to imprecision.

One additional study assessed subjective SL at the 15 mg dosage. Hindmarch[111] found no effect on the VAS scale rating for "ease of getting to sleep."

Three studies[114,115] evaluated the effects of temazepam 30 mg on subjective sleep latency from patient diaries. Fillingham[114] reported a reduction of SL of 40 min versus placebo. Heffron[115] found a 45 min reduction versus placebo. Hindmarch[111] reported a statistically significant effect on a VAS for "ease of getting to sleep" with temazepam 30 mg.

Tuk[116] found no difference between temazepam 20 mg and placebo in PSG SL. However, it is noteworthy that in this sample of "primary sleep onset insomnia" patients, both temazepam and placebo produced a reduction from baseline of approximately 53 min (to SL of about 24 min). Wilson[117] demonstrated a SL derived from actigraphy which was only 7 min less than that of placebo. However, of note, the end-of-treatment SL for temazepam (by actigraphy) was only 15 min, suggesting a possible floor effect for these results.

Three studies assessed subjective SL with temazepam 20 mg.[92,113,117] Cuanang[113] reported a reduction from baseline which was 34.2 min greater than placebo reduction. Voshaar[92] found end-of-treatment SL for temazepam 20 mg which was 29 min less than placebo. Similarly, Wilson[117] found subjective SL was 23 min less than placebo.

TOTAL SLEEP TIME: Two studies[110,112] were included in the meta-analysis for subjective TST at 15 mg. The analysis revealed a mean increase in TST of 64.4 min (CI: +8.1 to +120.8 min). Quality of evidence was moderate due to imprecision. No additional studies evaluated subjective TST at this dosage. Wu[112] reported a PSG TST of 99.1 min greater than placebo for 15 mg.

Two studies[114,115] reported subjective TST at the 30 mg dosage. Fillingim[114] demonstrated TST which was 53 min greater than placebo, while Heffron[115] noted a 54.6 min greater TST versus placebo. There were no investigations of objective TST for this dosage.

At the 20 mg dosage, three trials[92,113,117] reported subjective TST. Cuanang[113] found a 78 min greater TST increase from baseline than placebo. Voshaar[92] demonstrated a 46 min greater

TST than placebo at end-of-treatment. Wilson[117] also found an 18 min greater TST with temazepam 20 mg than with placebo. One study[117] assessed objective TST at 20 mg. Actigraphy revealed a 12 min greater TST at this dosage versus placebo.

Wake after sleep onset: Meta-analysis for WASO was not possible. One investigation[116] evaluated PSG WASO at the 20 mg dosage and reported WASO time which was 28.1 min less than placebo. Of note, the subjects in this study were described as exhibiting "sleep onset insomnia." At the same dosage, subjective WASO was 15 min less than placebo.[92] This was below the threshold for clinical significance.

Quality of sleep: Meta-analysis was conducted for sleep quality ratings from two studies[110,111] for temazepam 15 mg. The SMD was 0.25, below the range for clinical significance. However, it should be noted that the Hindmarch[111] study was underpowered to detect all but extremely large effects in that it only included 20 subjects. The quality of evidence was moderate due to imprecision.

Two studies found statistically significant improvement in sleep quality ratings for temazepam 30 mg.[114,115] Cuanang[113] reported statistically significant improvement for temazepam 20 mg on a quality rating comparing "better quality" to "same or worse quality."

Sleep efficiency: Meta-analysis was not achievable for SE at any dosage.

At 15 mg, Wu[112] found a PSG SE which was 13.3% greater than placebo (CI: +3.9 to +22.6%). Subjective SE was +14.1% versus placebo (CI: +5.8 to +22.3%). The quality of evidence for both was moderate due to imprecision. At 20 mg, Tuk[116] reported a +5.9% PSG SE versus placebo.

Number of awakenings: No meta-analysis of NOA was possible. One study[110] reported data for subjective NOA at the 15 mg dosage (−0.5 awakenings; CI: −1.29 to +0.29 awakenings). Quality of evidence was moderate due to imprecision.

Two studies[114,115] reported subjective NOA at 30 mg. They found −1.0 and −1.24 awakenings, respectively, compared to placebo.

Tuk[116] found no significant reduction in PSG NOA at 20 mg. One study[117] reported data for subjective NOA at 20 mg (−0.2 awakenings compared to placebo).

Overall quality of evidence: The overall quality of evidence in the meta-analytic data from the two available studies was moderate for temazepam 15 mg due to imprecision.

Harms: Limited data on adverse effects of temazepam 15 and 30 mg are available. Meta-analysis could not be performed. Glass[110] found no notable increase in adverse effects with temazepam 15 mg versus placebo and no significant effects were found on measures of daytime impairment. Cuanang[113] reported "no marked difference in adverse events," although temazepam 20 mg was associated with a modest increase in headache, blurred vision, depression and confusion. However, the frequency of these events was low overall. Heffron[115] found no difference in overall frequency of adverse events but noted that drowsiness, lethargy and vertigo were more commonly reported with temazepam 30 mg. There is some evidence that temazepam 30 mg is associated with daytime impairment on tests such as the Choice Reaction Time Test and Critical Flicker Fusion Test.[111]

In summary, the task force found that there was weak evidence of efficacy of temazepam in terms of therapeutic effects on sleep onset, total sleep time, awakenings, sleep efficiency, and possibly WASO with limited or no consistent evidence of adverse events in excess of placebo. However, there was also limited evidence for daytime impairment with temazepam 30 mg. Over, benefits were deemed to outweigh harms for temazepam 15 mg.

Patients' values and preferences: Based on its clinical judgement, the task force determined that a majority of patients would be likely to use both temazepam 15 mg and 30 mg compared to no treatment.

Melatonin agonists

Ramelteon for the Treatment of Chronic Insomnia

Recommendation 7: We suggest that clinicians use ramelteon as a treatment for sleep onset insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 8 mg doses of ramelteon.

Summary

Four RCTs investigated the use of ramelteon in the treatment of chronic primary insomnia.[118–121] The overall quality of evidence from these studies was downgraded to very low due to substantial heterogeneity across studies, imprecision and potential publication bias. The overall evidence for ramelteon 8 mg. was weakly in favor of its effectiveness for the treatment of sleep onset disturbance only. Meta-analysis of the three studies meeting inclusion criteria that reported objective (PSG) sleep latency demonstrated marginal reduction of sleep latency. The analysis revealed minimal increase in PSG-determined total sleep time which fell well below the defined threshold for clinical significance. Measures of sleep efficiency and sleep quality showed no clinically significant improvement. There was no evidence of significant difference from placebo for any adverse events, based on available side effect data. Although the evidence for efficacy is marginal, the benefits appear to be greater than the minimal potential harms. Based on clinical judgment, the task force determined that the majority of well-informed patients would use ramelteon over no treatment. This judgement is based on the evidence of improved sleep latency, coupled with its apparently low potential for adverse events.

See **Figures S31–S38**, **S76** and **S77** and **Table S14** in the supplemental material.

Discussion

Evidence from four RCTs which investigated the use of ramelteon in the treatment of chronic primary insomnia was included

in the statistical analysis.[118–121] Subjects in all studies demonstrated chronic primary insomnia with associated daytime complaints. All studies required mean objective LPS of > 20 min on two nights of PSG screening. All studies except Mayer and colleagues[119] also required mean objective WASO > 60 min. Kohsaka and colleagues[118] studied 65 chronic insomnia patients for two nights each at ramelteon doses of 4, 8, 16, and 32 mg. Roth and colleagues[120] studied 100 older adults (age > 65 years) with chronic primary insomnia. Subjects were administered two consecutive nights each of placebo, ramelteon 4 mg, and ramelteon 8 mg in a three-phase crossover protocol, with randomization of the treatment sequence and sustained washout time between each two night sequence.

Zammitt[121] studied the effects of nightly ramelteon in adults at dosages of 8 and 16 mg. PSG was conducted at baseline, and weeks 1, 3, and 5. The Mayer paper[119] reported on six-month nightly use of ramelteon in 451 adults with chronic insomnia from 46 multinational sites. Two nights of PSG were conducted in week 1 and at approximately one month intervals thereafter.

Sleep latency: The impact of 8 mg ramelteon on PSG-assessed SL was evaluated in three studies.[118,120,121] Objective sleep latency data in the study by Mayer and colleagues[119] were not adequate for meta-analysis and therefore could not be included.

Meta-analysis of the grouped evidence demonstrated marginal improvement in this critical outcome. However, the mean difference between the treatment and control groups was not clinically significant (−9.57 min; CI: −6.38 to −12.75 min). The confidence interval crossed the clinical significance threshold, and therefore the quality of evidence was downgraded for imprecision. It was downgraded further for the high degree of heterogeneity across studies (I^2 = 96%), and due to the risk of publication bias since all these studies were funded by industry. The resultant quality of evidence is very low.

Mean differences in objective sleep latency varied from −7.6 min to −13.1 min. Of note, the Roth investigation included exclusively older adults and found the smallest improvement in sleep latency. Subjective sleep latency from these investigations was comparable to objective latencies with mean difference (−11.44 min; CI: −3.31 to −19.56 min) falling below the clinical significance threshold.

Several additional papers which met inclusion criteria, but did not contain data suitable for this analysis, have addressed the efficacy and side effect profile of ramelteon.[122–126] The objective and subjective sleep latency from these results were consistent with the meta-analysis findings. This was likewise the case for sub-group analysis of subjects with primary sleep onset complaints.[124] A post-hoc analysis of the data from Zammitt by Mini and colleagues[123] found a significantly greater percentage of ramelteon 8 mg patients with > 50% reduction in sleep latency at week 1 (63.0% versus 39.7% for placebo), week 3 (63.0% versus 41.2%), and week 5 (65.9% versus 48.9%).

Total sleep time: All four studies included in the meta-analysis evaluated objective total sleep time for ramelteon 8 mg.[118–121] Although small improvements in TST were observed in some individual studies, ranging from 1.2 to 12.5 min longer, the meta-analysis reveals minimal increase (+6.58 min; CI: +1.36 to +11.80 min) which falls well below the threshold for clinical significance. The quality of evidence was downgraded to LOW due to the high degree of heterogeneity across studies, and due to the risk of publication bias since all these studies were funded by industry. Meta-analysis results of reported subjective TST were consistent with the objective finding (+5.7 min; CI: −7.65 to +19.04 min). Additional studies not included in meta-analysis supported these results.[122,125,126]

Wake after sleep onset: Meta-analysis of objective WASO from the two studies reporting adequate data[118,121] show a clinically insignificant increase (+3.5 min; CI: +2.77 to +4.23 min) in WASO for the ramelteon group, well below the significance threshold of 20 min. The quality of evidence was downgraded to moderate due to potential publication bias. One study not included in meta-analysis[122] found no difference in PSG WASO.

Zammitt and Mayer reported subjective WASO data for meta-analysis.[119,121] The ramelteon group demonstrated a clinically insignificant increase in WASO of 5.2 min (CI: −6.77 to +17.24 min). The quality of evidence was low due to heterogeneity and potential publication bias. The only additional study which assessed subjective WASO found no difference between placebo and ramelteon 8 mg.[126]

Quality of sleep: Sleep quality ratings showed virtually no difference from placebo in any of the studies assessed.[119–121] Meta-analysis suggests no difference between ramelteon and placebo, with a pooled mean difference of −0.04 points (CI: −0.13 to +0.05 points) on a 7-point Likert scale. The quality of evidence was downgraded to low due to heterogeneity and the risk of publication bias since all these studies were funded by industry. Additional studies which assessed subjective sleep quality found no difference between ramelteon and placebo groups.[122,125,126]

Sleep efficiency: Three studies reported sleep efficiency data included in meta-analysis.[118,120,121] Minimal improvements in sleep efficiency were reported (+1.93%; CI: +1.00 to +2.87%), falling well below the clinically significance threshold for objective sleep efficiency of 5%. The quality of evidence was low due to heterogeneity and potential publication bias. Additional studies did not report sleep efficiency data.

Number of awakenings: No meta-analysis for PSG number of awakenings was conducted as only one study reported adequate data for analysis.[119] This investigation found no clinically significant difference between ramelteon 8 mg and placebo (+0.1 awakenings; CI: +0.08 to +0.15 awakenings). The quality of evidence was moderate due to potential publication bias. Other studies which evaluated NOA reported no significant differences as well.[120,125,126]

In summary, these studies show very weak evidence for reduction of sleep latency at the recommended prescribed dosage (8 mg), with mean decrease of 9.57 min (CI: −6.38 to −12.75 min), and no consistent evidence of improvement in other objective or subjective parameters.

Overall quality of evidence: The overall quality of evidence in the meta-analytic data from these studies was

downgraded to very low for several reasons. Substantial heterogeneity across studies was noted for multiple outcomes. The data were also downgraded for imprecision, due to the relatively large confidence intervals, which cross the clinical significance thresholds for multiple outcomes. All of these studies were industry sponsored, resulting in further downgrading of evidence due to potential publication bias. The quality of evidence for individual outcomes ranged from moderate to very low; therefore, the overall quality of evidence was very low.

HARMS: Meta-analytic data on adverse effects showed a relatively low frequency of adverse effects overall and none which were significantly different than placebo. This analysis included headache, nausea, upper respiratory infection and nasopharyngitis. A single case of leukopenia, which was judged possibly related to medication, was noted in the Mayer study.[119] Both Zammitt[121] and Mayer[119] found no evidence of rebound insomnia or withdrawal effects following discontinuation (notably, the Mayer et al. study was based on six months of nightly use).

The studies not included in the meta-analysis found no indication of a significant difference in adverse events between ramelteon and placebo. Commonly reported adverse events in these studies included fatigue, headache, dizziness and somnolence.

Three studies assessed for next-day impairment associated with ramelteon. Roth and colleagues reported on next-day residual pharmacological effects of ramelteon in an older adult population.[120] Observations of DSST, immediate and delayed recall, subjective alertness, and concentration showed no significant residual as compared to placebo on any outcomes. Employing the same residual effect measures, Zammitt et al.[121] reported small but statistically significant impairment with ramelteon 8 mg. in immediate recall at week 3 only, delayed recall (week 1 only), level of alertness (week 5), and ability to concentrate (week 1). Mayer[119] found no consistent evidence of next-day impairment in alertness, recall, DSST or visual analogue scales of mood, energy, or cognition. Overall, the available data suggest no consistent evidence of next-day impairment associated with the use of ramelteon.

In summary, the task force found that there was weak evidence of efficacy in the treatment of sleep onset insomnia, with limited or no consistent evidence of adverse events in excess of placebo. Therefore, benefits were deemed to marginally outweigh harms.

PATIENTS' VALUES AND PREFERENCES: Based on its clinical judgement, the task force determined that in light of its efficacy for sleep onset and its relatively benign side effect profile, a majority of patients would be likely to use ramelteon compared to no treatment.

Heterocyclics

Doxepin for the Treatment of Chronic Insomnia

Recommendation 8: We suggest that clinicians use doxepin as a treatment for sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 3 mg and 6 mg doses of doxepin.

Summary

Four studies addressed the efficacy of doxepin 3 mg.[127–130] Four studies also investigated the 6 mg dosage.[128–131] The overall quality of evidence for both dosages was low due to potential publication bias and imprecision. The evidence suggests minimal improvement in SL but clinically significant improvements in WASO, TST and SE. The overall evidence was graded as weakly in favor of doxepin's efficacy in improving sleep maintenance.

Meta-analysis shows that PSG and patient-reported SL at 3 mg and PSG SL at 6 mg fell below the clinical significance threshold. Both PSG and subjective TST at 3 mg, as well as PSG TST at 6 mg, were above significance thresholds, although subjective TST at 6 mg fell short of this criterion. PSG data for reduction of WASO exceeded the clinical significance threshold at both dosages, although patient diary data for WASO at the 6 mg dosage fell below threshold, based on two studies. The SMD in sleep quality for doxepin 3 mg suggests moderate improvement, while the SMD for the 6 mg dosage suggests mild improvement. The objective SE for both dosages exceeded the clinical significance level, while objective NOA fell short.

Meta-analysis of side effects included headache, diarrhea, somnolence and upper respiratory infection at 3 mg, and headache and somnolence at the 6 mg dose. Results suggest mild increase in somnolence at 6 mg. Given the demonstrated improvements in WASO, TST and SE, with limited adverse effects, the task force judged the benefits to outweigh the harms. The clinical judgement of the task force was that the majority of well-informed patients would use doxepin over no treatment. This judgement is based on the evidence for clinically significant improvement in WASO, TST and SE.

See **Figures S39–S53**, **S78–S83** and **Tables S15** and **S16** in the supplemental material.

Discussion

Five studies investigated the effects of doxepin at 3 mg and/or 6 mg.[127–131] Krystal[127] conducted a 12-week RCT of nightly doxepin 1 and 3 mg versus placebo in 240 elderly (> 65 years) subjects with predominant sleep maintenance insomnia. Subjects were randomized to one of three treatment groups. Outcome variables included both PSG and sleep diary data. Krystal[128] investigated doxepin 3 mg and 6 mg in a five week trial which included 221 adults with sleep onset and maintenance insomnia who were randomized to one of the two doxepin doses or placebo. PSG data and sleep diaries were included. Roth[129] employed a crossover design with randomized assignment to one of four treatment sequences which consisted of two nights each of doxepin 1 mg, doxepin 3 mg, doxepin 6 mg and placebo, with intervening washout. PSG and sleep diary data were collected. The study included 67 adults who met both baseline PSG-defined sleep onset and maintenance criteria. Scharf[130] employed the identical crossover design and dosages in 76 elderly insomnia subjects. Lankford[131] reported data on a four week nightly trial of doxepin 6 mg or placebo in 254 elderly

subjects with sleep onset and sleep maintenance insomnia. Outcome variables were patient-reported and clinician rated.

Hajak[132] also conducted a RCT of doxepin, but the dosages (25–50 mg) were significantly higher that FDA-approved hypnotic dosages. For this reason, this study was not included in the current analysis.

Sleep latency: Four studies[127–130] reported PSG SL data for the 3 mg dosage. The mean difference from placebo (−2.30 min; CI: −6.22 to +1.62 min) was below the defined significance threshold. Evidence quality was moderate due to potential publication bias. Likewise, patient-reported SL[127,130] did not meet clinical significance (−9.35 min; CI: −21.89 to +3.19 min). Quality was low due to imprecision and potential publication bias. Three studies included adequate data for meta-analysis at the 6 mg dosage,[128–130] showing a mean difference for objective SL of −5.29 min (CI: −1.34 to −9.25 min) with moderate quality of evidence due to publication bias. No sleep diary data were available for meta-analysis of SL at this dosage.

Total sleep time: Four investigations[127–130] reported PSG data for TST at 3 mg. The analysis reveals a clinically significant increase in TST at this dosage (+26.14 min; CI: +18.49 to +33.79 min). Quality was low due to imprecision and potential publication bias. Subjective reports for 3 mg[127,130] were also in the range of clinical significance (+43.57 min; CI: +5.16 to +81.98 min) with very low quality of evidence due to heterogeneity, imprecision and potential publication bias. At the 6 mg dosage, PSG-determined TST,[128–130] also met the clinical significance criterion (+32.27 min; CI: +24.24 to +40.30 min) with moderate quality of evidence due to potential publication bias. However, subjective TST at this dosage[130,131] fell short of significance (+18.84 min; CI: −1.65 to +39.34 min) with LOW quality of evidence due to imprecision and potential publication bias.

Wake after sleep onset: WASO was considered a key outcome variable in all of the doxepin studies noted. The PSG data for 3 mg doxepin showed a clinically significant mean difference from placebo of −22.17 min (CI: −14.72 to −29.62 min), based on four trials.[127–130] Quality of evidence was low due to imprecision and potential publication bias. Only one study reported subjective WASO, with a reduction of 20.0 min versus placebo. Quality of these data was low due to imprecision and potential publication bias. At 6 mg, PSG WASO showed a clinically significant reduction of 23.14 min (CI: −16.36 to −30.34 min)[128–130] with LOW quality of evidence due to imprecision and potential publication bias. Patient diary results did not meet clinical significance (−14.39 min; CI: −3.93 to −24.86 min)[130,131] with moderate quality of evidence due to potential publication bias.

Quality of sleep: Quality of sleep ratings for the 3 mg dosage suggest substantial improvement (SMD: +0.57; CI: +0.26 to 0.88 SMD) with low quality of evidence,[127,130] due to imprecision and potential publication bias. More modest improvement was noted at 6 mg (SMD +0.28; CI +0.06 to 0.49 SMD)[130,131] with moderate quality of evidence due to potential publication bias.

Sleep efficiency: PSG SE was reported in three studies for the 3 mg dosage.[127,129,130] Evidence quality was low due to imprecision and potential publication bias. The improvement in SE was clinically significant at +6.78% (CI: +4.50 to 9.07%). SE at the 6 mg dose, based on two investigations[129,130] was also significantly improved (+7.06%; CI: +5.12 to 9.01%) with moderate quality of evidence due to potential publication bias.

Number of awakenings: PSG-determined NOA was mildly *increased* (+0.53 awakenings; CI: −0.37 to +1.42 awakenings) for 3 mg[127,129,130] and the 6 mg dose (+0.44 awakenings; CI: −0.57 to +1.44 awakenings), with moderate quality for both, due to potential publication bias.

Overall quality of evidence: The quality of evidence in the meta-analytic data for the majority of variables was moderate to low due to industry sponsorship and, in some cases, imprecision (due to relatively large confidence intervals for numerous variables that cross clinical significance thresholds). Quality was further downgraded to very low for subjective TST at 6 mg as a result of the above factors plus heterogeneity of data. As a result, the overall quality of evidence for the doxepin data is considered very low.

Harms: Meta-analysis was available for both the 3 mg[127–129] and 6 mg[128,129,131] dosages and revealed no increase in headache frequency with doxepin. Somnolence showed no significant increase versus placebo (+0.01 risk difference) at the 3 mg level[127–129] and a small increased risk at 6 mg (+0.04 risk difference).[128,129,131] Data were also available for meta-analysis of risk for diarrhea and upper respiratory infection. Neither showed significantly greater risk than placebo. With respect to next-day residual effects, no difference was observed between doxepin 3 mg or 6 mg and placebo on DSST, Symbol Copying Test, or visual analogue scales for morning sleepiness.[127–130]

In summary, the task force found weak evidence for efficacy in the treatment of sleep maintenance insomnia, with minimal evidence of adverse events in excess of placebo. Therefore, benefits were deemed to be greater than harms.

Patients' values and preferences: Based on its clinical judgement, the task force determined that in light of the data supporting efficacy for reducing WASO, and improving TST, SE and sleep quality, a majority of patients would be likely to use doxepin compared to no treatment.

Trazodone for the Treatment of Chronic Insomnia

Recommendation 9: We suggest that clinicians not use trazodone as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on one trial of a 50 mg dose of trazodone.

See **Table S17** in the supplemental material.

Summary

A single study[78] of trazodone 50 mg met inclusion criteria; therefore, no meta-analysis is available. The overall quality of evidence for this study was moderate due to potential publication bias. The patient-reported data from this study demonstrated a modest reduction in SL which fell below the threshold for clinical significance. Likewise, the moderate increase in TST and the small reduction in WASO did not reach the clinical threshold criteria. Quality of sleep was insignificantly improved and reduction in NOAs fell just below clinical significance. In summary, none of the sleep outcome variables improved to a clinically significant degree.

No meta-analysis of harms was possible. Given the absence of demonstrated efficacy on numerous critical outcome variables, coupled with limited evidence regarding harms, the task force judged the harms to potentially outweigh the benefits. Based on its clinical judgement, the task force determined that, despite the absence of significant efficacy for trazodone 50 mg and the paucity of information regarding harms, the majority of patients would be likely to use trazodone compared to no treatment.

Discussion

Walsh[78] investigated the efficacy of trazodone 50 mg versus zolpidem 10 mg and placebo. The final sample for the trazodone and placebo groups included 187 adults with sleep onset insomnia. Subjects were administered either trazodone or placebo in double-blind fashion for 14 consecutive nights. All data were patient-reported.

SLEEP LATENCY: Subjective SL was reduced by 10.2 min (CI: −8.95 to −11.44 min). This falls short of the clinical significance threshold. The quality of evidence was moderate due to potential publication bias.

TOTAL SLEEP TIME: Sleep diary TST was increased by a clinically insignificant 21.8 min (CI: +20.10 to +23.49 min). The quality of evidence was moderate due to potential publication bias.

WAKE AFTER SLEEP ONSET: Sleep diary WASO was reduced by 7.7 min (CI: −8.89 to −6.5 min), falling below the threshold. The quality of evidence was moderate due to potential publication bias.

QUALITY OF SLEEP: On a 4-point scale (1 = excellent, 4 = poor) sleep quality was not significantly improved versus placebo (−0.13 points; CI: −0.11 to −0.14 points). The quality of evidence was moderate due to potential publication bias.

NUMBER OF AWAKENINGS: This outcome was reduced by 0.4 (CI: −0.37 to −0.42 awakenings) compared to placebo, less than the 0.5 subjective awakening threshold. The quality of evidence was moderate due to potential publication bias.

OVERALL QUALITY OF EVIDENCE: The overall quality of evidence for this study was moderate.

HARMS: There was no meta-analysis of harms. In the Walsh[78] paper, the trazodone group experienced significantly more side effects than the placebo group. Chief among these were headache (trazodone 30%; placebo 19%) and somnolence (trazodone 23%; placebo 8%). In all, 75% of trazodone subjects reported some adverse event(s), compared to 65.4% of subjects who received placebo.

PATIENTS' VALUES AND PREFERENCES: Based on its clinical judgement, the task force determined that, despite the absence of significant efficacy for trazodone 50 mg and the paucity of information regarding harms, the majority of patients would be likely to use trazodone compared to no treatment. This is based on the perception of trazodone as a "safer" sleep-promoting agent by many physicians and the resulting recommendations and prescribing practices of those physicians.

Anticonvulsants

Tiagabine for the Treatment of Primary Insomnia

Recommendation 10: We suggest that clinicians not use tiagabine as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 4 mg doses of tiagabine.

Summary

Three studies addressed the efficacy of tiagabine 4 mg.[133–135] The overall quality of evidence was very low due to potential publication bias, heterogeneity, and imprecision. Meta-analyses were conducted for SL (PSG and subjective), TST (PSG and subjective), WASO (PSG and subjective), sleep quality, SE (PSG), and NOA (PSG and subjective). These analyses revealed that both objective and subjective measures of sleep latency fell below the threshold for clinical significance. Measures of TST showed minimal change (PSG) and mild to moderate *reduction* (sleep diary). WASO data demonstrated no clinically significant change on either metric. Meta-analysis of SMD for sleep quality suggested improvement which fell below the clinical significance threshold. Neither objective nor subjective NOAs were reduced by clinically significant levels, while PSG SE was minimally *reduced*.

Meta-analysis of adverse effects showed no difference between tiagabine and placebo on headache or nausea. Given the absence of demonstrated efficacy on numerous critical outcome variables (with slight trending toward mild worsening on some outcomes), coupled with limited evidence regarding harms, the task force judged the harms to potentially outweigh the benefits.

It was determined by clinical judgement of the task force that the majority of well-informed patients would not use tiagabine over no treatment. This judgement is based on the lack of evidence for efficacy and the limited systematic information regarding adverse effects.

See **Figures S55–S64**, **S84** and **S85** and **Tables S18–S20** in the supplemental material.

Discussion

Three studies were included in the meta-analyses of tiagabine.[133–135] Roth[133] studied 207 elderly primary insomnia patients (65–85 years) with difficulty initiating and maintaining sleep who received tiagabine 2, 4, 6, or 8 mg or placebo on two consecutive study nights with PSG recordings in a parallel group design. Walsh[134] similarly evaluated 232 adults with chronic sleep-onset and maintenance insomnia. Tiagabine 4, 6, 8, or 10 mg or placebo was administered on two consecutive nights with PSG. Walsh[135] conducted a crossover study of 58 adults (age 35–64) with chronic sleep onset and maintenance problems. Subjects received 4, 8, 12, and 16 mg and placebo for two consecutive nights of sleep recording. Medication-free washout periods between doses ranged from 5–12 nights.

Sleep latency: The meta-analysis for SL included three studies.[133–135] PSG SL data showed a small *increase* in SL (+3.65 min; CI: −8.00 to +15.31 min) with very low quality of evidence due to heterogeneity, imprecision and potential publication bias. The subjective data[133,135] showed a moderate *increase* in SL (+13.31 min; CI: +7.54 to 19.37 min). Quality of evidence was moderate due to potential publication bias.

Total sleep time: Objective data for TST[133–135] demonstrated a minimal *reduction* in TST (−1.21 min; CI: −7.44 to +5.02 min) with LOW quality evidence due to heterogeneity and potential publication bias. Patient-reported TST[133,135] was *reduced* by 19.95 min (CI: −25.35 to −14.54 min) with moderate quality of evidence due to potential publication bias. Neither subjective nor objective findings met clinical significance.

Wake after sleep onset: The PSG WASO analysis[133–135] revealed essentially no difference from placebo (−0.56 min; CI: −6.77 to +5.65 min). Quality of evidence was low due to heterogeneity and potential publication bias. Sleep diary data[133,135] indicated a small, clinically insignificant *increase* (+4.29 min; CI: −0.22 lower to +8.79 min) with moderate quality of evidence due to potential publication bias.

Quality of sleep: The meta-analysis for QOS[133,135] resulted in a SMD of +0.48 (CI: −0.5 to +1.46 SMD), which falls below the level of clinical significance. Quality of evidence was very low due to heterogeneity, imprecision and potential publication bias.

Sleep efficiency: The objective sleep efficiency was *reduced* (−0.53%; CI: −0.02 to −1.05%). Quality of evidence was moderate due to potential publication bias.

Number of awakenings: The PSG NOAs were mildly *increased* (+0.5 awakenings; CI: −1.29 to +2.29 awakenings). The subjective NOA was minimally reduced at −0.21 awakenings (CI: −0.9 to +0.48 awakenings), falling below the threshold for clinical significance. Level of evidence was low for both measures due to imprecision and potential publication bias

Overall quality of evidence: The overall quality of evidence for the meta-analytic data was very low due to significant heterogeneity, imprecision and potential bias (industry sponsorship) for some critical outcomes.

Harms: Meta-analysis was possible for two adverse effects (headache and nausea). Neither showed any significant difference from placebo. None of the three studies found a significant difference from placebo on morning-after DSST or visual analogue scales for sleepiness/alertness at the 4 mg dose.

Patients' values and preferences: Based on its clinical judgement, the task force determined that in light of the absence of significant efficacy at this dose and the paucity of information regarding harms, the majority of patients would not be likely to use tiagabine compared to no treatment.

Over-the-counter preparations

Diphenhydramine for the Treatment of Primary Insomnia

Recommendation 11: We suggest that clinicians not use diphenhydramine as a treatment for sleep onset and sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 50 mg doses of diphenhydramine.

Summary

Two RCTs evaluated diphenhydramine 50 mg for the treatment of chronic primary insomnia.[110,136] The overall quality of evidence was downgraded to low due to imprecision and risk of publication bias. The overall evidence for diphenhydramine 50 mg was weakly against its effectiveness for improving sleep onset and TST. The mean reduction in patient-reported sleep latency versus placebo fell below the level of clinically significant improvement. The same studies found a small increase in TST which also fell below the threshold for clinical significance. The single paper[136] which included PSG-determined SL and TST showed outcomes which also fell below clinical significance thresholds. None of the other objective or patient-reported outcome variables reached clinical significance thresholds. In addition, one paper meeting inclusion criteria[137] but not including suitable data for meta-analysis evaluated diphenhydramine 50 mg in mild to moderate insomnia patients.

No meta-analysis was possible for side effects. Since no systematic data addressing harms is available, it is difficult to make a clear determination regarding benefits versus harms. However, in light of the absence of clear benefits, the task force judged the benefits and harms to be approximately equal. It was determined by clinical judgement of the task force that the majority of well-informed patients would not use diphenhydramine over no treatment. This judgement is based on the absence of evidence for clinically significant improvement.

See **Figures S65** and **S66** and **Table S21** in the supplemental material.

Discussion

Two studies of diphenhydramine 50 mg included adequate data for meta-analysis. Glass[110] studied 25 elderly subjects (mean age = 73.9 years) with insomnia. Enrollees received diphenhydramine, temazepam 15 mg and placebo in a crossover design with two weeks of nightly use for each intervention, followed by washout. Primary outcomes measures were sleep variables recorded in patient diaries. Morin[136] compared diphenhydramine (14 nights, followed by 14 nights of placebo) to a valerian-hops preparation (28 nights) and placebo (28 nights) in a total population of 184 adults with occasional insomnia (2–4 nights/week with SL > 30 min or WASO > 30 min). Patients were randomly assigned to the intervention groups and PSG and patient-reported data were collected. A third study,[137] not included in meta-analysis, assessed mild to moderate insomnia patients in family practice settings. Participants received diphenhydramine 50 mg and placebo for one week each in crossover fashion, without intervening washout. Outcome assessment was based on patient-completed sleep questionnaires.

Sleep latency: The single study employing PSG[136] found a 7.89 min reduction in SL (CI: −17.40 to +1.62 min). This fell below the significance threshold. Quality of evidence was low due to imprecision and potential publication bias. Two studies[110,136] met requirements for meta-analysis of subjective SL. This revealed a mean difference from placebo of −2.47 min (CI: −8.17 to +3.23 min). The Rickels study[137] found statistically significant improvement in SL with diphenhydramine using a 0–4 patient-rating scale, but no specific quantitative data regarding actual SL times were included.

Total sleep time: Morin[136] reported a PSG TST increase of 12.37 min (CI: −13.38 to +38.12 min). This fell below the significance threshold of 20 min. Quality of evidence was low due to imprecision and potential publication bias. Meta-analysis of the two studies demonstrated a 17.86 min increase (CI: −3.79 to + 39.51 min) in subjective TST versus placebo. The Rickels study[137] found "statistically significant improvement" in patient-reported TST but, as noted above, it is unclear to what extent this represented clinically significant improvement.

Wake after sleep onset: No data pertaining to wake after sleep onset were available.

Quality of sleep: Glass[110] found minimal difference in sleep quality between diphenhydramine and placebo (mean difference of +0.1 SD; CI: −0.45 to +0.65 SD). Quality of evidence was downgraded to moderate due to potential publication bias. Rickels[137] reported statistically significant improvement in sleep quality.

Sleep efficiency: The objective sleep efficiency data from the single study reporting PSG analysis[136] found no clinically significant improvement (+2.59%; CI: −3.25 to +8.43%). In this same study, subjective SE also fell below the threshold (+4.61%; CI: +1.33 to +7.88%).

Number of awakenings: The change in subjective number of awakenings (−0.3 awakenings; CI: −1.03 to +0.43 awakenings) was not clinically significant.[110]

Overall quality of evidence: The overall quality of evidence in the meta-analytic data from these studies was downgraded to low for imprecision, due to confidence intervals which crossed the clinical significance thresholds for subjective TST, a critical outcome. These studies were industry sponsored, resulting in further downgrading of evidence due to potential publication bias. The quality of evidence for individual critical outcomes ranged from moderate to low, therefore the overall quality of evidence was low.

Harms: No meta-analysis of adverse effects was possible. Neither Morin[136] nor Glass[110] found significant differences between diphenhydramine and placebo in adverse events. Rickels[137] reported higher numerical rates of drowsiness, dizziness, and grogginess with diphenhydramine but no statistical analysis was conducted.

Morin[136] found no substantial rebound effects following discontinuation of diphenhydramine. Glass[110] noted minimal differences between diphenhydramine and placebo in the number of subjects experiencing rebound for at least one sleep outcome variable. Glass[110] found no difference in morning-after DSST or Manual Tracking Task (MTT) between interventions.

In summary, the task force found that there was weak evidence demonstrating an *absence* of efficacy in the treatment of sleep onset insomnia, with minimal evidence of adverse events in excess of placebo. Therefore, benefits were deemed approximately equal to harms.

Patients' values and preferences: Based on its clinical judgement, the task force determined that, in light of the paucity of data supporting efficacy for sleep onset and maintenance, a majority of patients would not be likely to use diphenhydramine compared to no treatment.

Melatonin for the Treatment of Primary Insomnia

Recommendation 12: We suggest that clinicians not use melatonin as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 2 mg doses of melatonin.

Summary

Three studies addressed the efficacy of melatonin 2 mg.[138–140] These investigations included only older adults (> 55 years). The overall quality of evidence was very low due to potential publication bias, heterogeneity, and imprecision. Meta-analysis was only achievable for sleep quality. This indicated a SMD of +0.21 (CI: −0.36 to +0.77 SMD), which was not clinically significant. The minimal overall evidence available was weakly against melatonin's efficacy in improving sleep onset, maintenance, or quality.

No adequate data for meta-analysis of adverse effects was available. Given the lack of evidence for efficacy in treating

insomnia, and the unavailability of systematic data on side effects, the task force judged the benefits to be approximately equal to harms. It was determined by the task force that the majority of well-informed patients would use melatonin over no treatment. This is based on its availability and the widespread perception of melatonin as a benign sleep aid.

See **Figure S67** and **Table S22** in the supplemental material.

Discussion

Three studies included adequate data for melatonin meta-analysis.[138–140] Lemoine[138] studied 170 older adults (age > 55 years) with primary insomnia. Subjects received either prolonged release melatonin (PRM) 2 mg or placebo nightly for 3 weeks. Outcome data was patient-reported. Luthringer[139] similarly studied 40 older adults (age > 55 years) who received PRM 2 mg or placebo for 3 weeks. Outcomes included both PSG and subjective data. Finally, Wade[140] evaluated 354 patients of the same age group with PRM 2 mg or placebo nightly for 3 weeks. Outcome data was patient-reported.

In addition, seven trials which met inclusion criteria but did not include adequate data for meta-analysis were identified.[141–147] These investigations employed various dosages and combinations with other agents, rendering meaningful comparisons to the 2 mg RCTs impossible. Pertinent features of these studies are included within each outcome section.

Haimov[143] conducted a randomized crossover study of elderly adults with insomnia consisting of one week on each of three interventions (2 mg sustained-release melatonin, 2 mg fast-release melatonin or placebo) with intervening washout, followed by a 2-month extension of 1 mg slow-release melatonin. Data were derived from actigraphy. Zhdanova[147] evaluated three dosages of melatonin (0.1, 0.3, and 3 mg) versus placebo in a randomized crossover study of 30 elderly (> 50 years) adults (15 normal sleepers and 15 insomnia subjects with reduced SE). Subjects received each dosage or placebo for one week with intervening washout. Wade[146] administered prolonged-release melatonin 2 mg or placebo to adults with primary insomnia for 3 weeks, after which the melatonin group continued for 26 weeks, while the placebo group was re-randomized to melatonin or placebo (1:1). Sleep outcome variables (from sleep diary) were analyzed according by age group as well as by melatonin deficiency status. Baskett[141] conducted a randomized controlled crossover study of healthy elderly with sleep maintenance problems. Subjects received 5 mg melatonin or placebo for four weeks with intervening washout.

Sleep latency: Meta-analysis was not possible for sleep latency. Luthringer[139] reported a PSG SL reduction of 8.9 min (CI: −2.35 to −15.45 min), which falls below clinical significance (prolonged release 2 mg). The quality of evidence was low due to imprecision and potential publication bias.

In the Haimov[143] investigation, fast-release melatonin produced significantly shorter SL than placebo at one week. At 2 months, sustained release 1 mg resulted in significantly shorter SL than placebo. Zhadanova[147] reported no significant improvement in PSG SL at any dosage.

Wade[146] found that the melatonin deficient group (including all ages) showed no improvement with melatonin versus placebo on SL at three weeks. However, the elder group (65–80 years) showed significant reduction of SL with melatonin, regardless of melatonin deficiency status (SL: −19.1 min; placebo −1.7 min). This improvement held at 19 weeks for the elder group (melatonin: −25.9 min; placebo: −8.3 min). Wade[145] subsequently re-analyzed these data and reported that the significant improvement in SL held when the age range for the "elderly" group was expanded to 55–80 years, but not lower. Baskett[141] found no improvement in SL (as measured by sleep diary) with melatonin 5 mg.

Total sleep time: There were inadequate data for meta-analysis of TST. Luthringer[139] found an increase of 2.2 min versus placebo (CI: −19.13 to +23.53 min) in objective TST with melatonin 2 mg. The quality of evidence was very low due to significant imprecision of the data, and potential publication bias.

Zhdanova[147] observed no increase in objective TST at any dosage. Wade[146] reported no improvement in patient-reported TST in the low melatonin secretor population (regardless of age) at 3 weeks but observed a small improvement (estimated difference: +13.1 min) at 29 weeks. Analysis of the elderly population revealed no significant improvement in TST at any point. Baskett[141] reported no improvement at the 5 mg dose as measured by sleep diary.

Wake after sleep onset: No meta-analysis was possible for WASO. Luthringer[139] found a small *increase* in WASO (+8.5 min: CI: −11.75 to +28.75 min) in the prolonged release melatonin 2 mg group. The quality of evidence was very low due to significant imprecision of the data, and potential publication bias.

Quality of sleep: The meta-analysis of QOS demonstrated a small improvement in quality of sleep (+0.21 SMD: CI: −0.36 to +0.77 SMD), which fell below the threshold for clinical significance. The quality of evidence was very low due to heterogeneity, imprecision and potential publication bias.

Baskett[141] found no improvement in quality of sleep with 5 mg melatonin. Wade[146] reported no improvement with prolonged-release melatonin at 3 weeks and 29 weeks in the low excretor and elderly groups.

Sleep efficiency: There were not adequate data for meta-analysis of melatonin SE.

Haimov[143] reported small to moderate increases in actigraphic SE versus placebo (placebo: 77.4%; fast-release 2 mg/1 week: 78.8%; sustained release 2 mg/1 week: 80.4%; sustained release 1 mg/2 months: 84.3%). Both of the sustained release dosages and durations were statistically significantly different from placebo. Zhdanova[147] also reported significant improvement in PSG SE versus placebo in the multiple dose crossover study (placebo: 78%; melatonin 0.1 mg: 84%; 0.3 mg 88%; 3 mg: 84%). Baskett[141] found no difference between placebo and melatonin 5 mg for subjective SE.

Number of awakenings: Insufficient data precluded meta-analysis of NOA. Luthringer[139] found an *increased* (+1.4

awakenings; CI: −4.59 to +7.39 awakenings) NOA with melatonin, as measured by PSG. The quality of evidence was very low due to significant imprecision of the data and potential publication bias.

Zhdanova[147] and Baskett[141] reported no difference in NOA between melatonin and placebo by PSG or patient diary, respectively.

Overall quality of evidence: The overall quality of evidence in the single outcome meta-analytic data from these studies was downgraded to very low due to heterogeneity, imprecision, and industry sponsorship, resulting in potential publication bias.

Harms: Meta-analysis for adverse events was not possible. Of the included investigations, none reported clinically significant differences in adverse events between melatonin and placebo for any dosage or duration.[138–140,146] With one possible exception, no rebound or withdrawal effects were reported.[138,139,146] Haimov[143] found marginally significant difference in SE between the active phase for two month, 1 mg sustained-release melatonin and the withdrawal period.

In summary, the task force found that there was weak evidence against clinically significant efficacy in the treatment of sleep onset insomnia, with little systematic evidence regarding harms. However, mixed evidence suggests possible improvement in SL in an elderly population. Therefore, benefits were deemed to be approximately equal to harms.

Patients' values and preferences: Based on clinical judgement, the task force determined that despite the paucity of meta-analytic data, equivocal data regarding efficacy for sleep-onset insomnia, and absence of data regarding sleep maintenance, a majority of informed patients would be likely to use melatonin compared to no treatment. As previously noted, this is based on its availability and the widespread perception of melatonin as a benign sleep aid.

L-tryptophan for the Treatment of Primary Insomnia

Recommendation 13: We suggest that clinicians not use tryptophan as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of 250 mg doses of tryptophan.

Summary

Only one study[148] contained adequate data, so meta-analysis was not possible. The quality of evidence for the critical outcomes was high. This study, consisting of patient-reported data, showed a modest decline in TST, which was not clinically significant. WASO was decreased slightly, while sleep quality was mildly increased; neither met thresholds for clinical significance. Sleep efficiency was insignificantly decreased.

No meta-analysis of harms was possible. Given the absence of demonstrated efficacy on numerous critical outcome variables, coupled with limited evidence regarding harms, the task force judged the harms to potentially outweigh the benefits. Based on its clinical judgement, the task force determined that, despite the absence of significant efficacy for tryptophan 250 mg and the absence of information regarding harms, the majority of patients would be likely to use tryptophan compared to no treatment.

See **Table S23** in the supplemental material.

Discussion

Hudson[148] investigated the effects of food source tryptophan (250 mg), pharmacological tryptophan 250 mg, both with carbohydrate, versus carbohydrate alone. Subjects (n = 31) received one of the three interventions for one week. Outcome data consisted of sleep diaries.

Two additional papers met inclusion criteria, but used much higher dosages. Hartmann[149] compared tryptophan 1 g to secobarbital, flurazepam, and placebo in a one week trial. Tryptophan and placebo groups included 52 subjects with chronic insomnia. Data were patient-reported. Spinweber[150] studied 20 young men with sleep onset insomnia. Following placebo run-in, ten subjects received tryptophan 3 g and ten received placebo for six nights, with PSG recordings nightly.

Sleep latency: The Hudson[148] study did not report sleep latency data.

Spinweber[150] noted improvement in PSG sleep latency only on nights 4–6 of administration (11.2 min lower than placebo for this period). Hartmann[149] found no difference in subjective sleep latency between tryptophan and placebo during active treatment.

Total sleep time: Hudson[148] reported a moderate *reduction* in subjective TST (−20 min; CI: −31.29 to −8.7 min). The quality of evidence was moderate due to imprecision. Other investigations did not report TST data.

Wake after sleep onset: Hudson[148] noted a small reduction in subjective WASO (−9.7 min: CI −15.21 to −4.18 min), that did not meet clinical significance. The quality of evidence was high.

Quality of sleep: On a 3-point scale (1 = low, 3 = high) sleep quality was increased (+0.3 points: CI +0.22 to +0.37 points) in the Hudson study.[148] The quality of evidence was high. Hartmann[149] found no significant difference between tryptophan and placebo on a measure of "How well I slept."

Sleep efficiency: Sleep efficiency was not reported by any study.

Number of awakenings: NOA was not reported by any study.

Overall quality of evidence: The overall quality of evidence for this the critical outcomes was high.

Harms: There was no meta-analysis of harms. None of the papers reported systematic information regarding adverse effects associated with tryptophan.

PATIENTS' VALUES AND PREFERENCES: Based on clinical judgement, the task force determined that, despite the absence of significant efficacy for tryptophan 250 mg and the absence of information regarding harms, the majority of patients would be likely to use tryptophan compared to no treatment.

Valerian for the Treatment of Primary Insomnia

Recommendation 14: We suggest that clinicians not use valerian as a treatment for sleep onset or sleep maintenance insomnia (versus no treatment) in adults. [WEAK]

Remarks: This recommendation is based on trials of variable dosages of valerian and valerian-hops combination.

Summary

Morin[136] evaluated a combination of valerian (374 mg native extract) and hops (83.8 native extract). The overall quality of evidence for these data was low due to imprecision and potential publication bias. PSG sleep latency was reduced to a degree that fell below the clinical significance threshold. Other measures, including subjective SL, as well as PSG and patient-reported TST and SE were improved, but did not meet clinical significance thresholds.

No meta-analysis of harms was possible. Given the absence of demonstrated efficacy on critical outcome variables (with the possible exception of marginally improved PSG SL), coupled with limited evidence regarding harms, the task force judged the harms to be roughly equal to the benefits. Based on its clinical judgement, the task force determined that, given the lack of efficacy for valerian (with the possible exception of small improvements in SL) and the limited information regarding harms, the majority of patients would not be likely to use valerian compared to no treatment.

See **Table S24** in the supplemental material.

Discussion

Morin[136] investigated the effects of a valerian-hops combination in dosages noted above. This combination was compared to diphenhydramine and placebo. Subjects with mild difficulty initiating or maintaining sleep were randomized to one of the three interventions (valerian-hops n = 59; diphenhydramine n = 60; placebo n = 65) with nightly administration for 28 days. A subset (valerian n = 22; placebo n = 26) underwent PSG at baseline and at the end of weeks one and two.

One additional paper[151] met inclusion criteria, but employed a higher dosage. Oxman conducted a randomized trial involving 405 adults of all ages with insomnia. Subjects were randomized to two-week, nightly administration of valerian (3,600 mg) or placebo. Outcomes were patient-reported and captured as ranges, therefore the data were not usable for meta-analysis.

SLEEP LATENCY: Morin[136] found a reduction in PSG SL of 9.29 min (CI: −0.27 to −18.3 min). This approached clinical significance. The quality of evidence was LOW due to imprecision and potential publication bias. Subjective SL, however, was *increased* by +3.77 min (CI: −4.47 to +12.01 min), with moderate quality of evidence due to potential publication bias. Oxman[151] found no statistically significant improvement in SL.

TOTAL SLEEP TIME: In the Morin[136] study, PSG TST was increased, although not to a clinically significant degree (+10.96 min; CI: −21.67 to +43.59 min) (very low quality of evidence). Patient-reported TST was higher (+3.12 min; CI: - 22.08 to +28.32) with moderate quality of evidence. Oxman[151] found no significant improvement in subjective TST.

WAKE AFTER SLEEP ONSET: WASO data were not reported in any study.

QUALITY OF SLEEP: Morin[136] did not report quality of sleep data and Oxman[151] found no statistically significant difference versus placebo in the percentage of patients meeting the defined sleep quality improvement criterion (valerian 28.7%; placebo 21.2%; difference +7.5% [95% CI:15.9 to 20.9%]).

SLEEP EFFICIENCY: Minimal increases in objective (+0.96%; CI: −5.02 to +6.94%) and subjective (1.85%; CI: −1.9 to +5.6%) SE were noted by Morin.[136] Both outcomes were downgraded due to potential publication bias, while PSG data was downgraded further due to significant imprecision.

NUMBER OF AWAKENINGS: Oxman[151] observed a statistically significant reduction in average change scores for NOA with valerian.

OVERALL QUALITY OF EVIDENCE: Quality of evidence for all outcomes ranged from very low to moderate. The only critical outcome for which adequate data was reported demonstrated low quality evidence, therefore the overall quality of evidence was low.

HARMS: Morin[136] observed no difference between valerian-hops and placebo with respect to frequency of adverse events. No serious adverse events were noted. Likewise, Oxman[151] found no increase in adverse events at the higher valerian dose compared to placebo.

PATIENTS' VALUES AND PREFERENCES: Based on its clinical judgement, the task force determined that, given the lack of efficacy for valerian (with the possible exception of small improvements in SL) and the limited information regarding harms, the majority of patients would not be likely to use valerian compared to no treatment.

LITERATURE REVIEWS

The following section contains literature reviews of drugs for which clinical practice recommendations were not possible, due to inadequate data for statistical analyses.

Estazolam

Summary

Three studies evaluated the efficacy of estazolam[152–154] using similar patient sleep questionnaires, but none of the data were suitable for meta-analysis. Likewise, it was not possible

to evaluate these data with respect to the established clinical significance thresholds. Therefore, no recommendations regarding efficacy of estazolam are possible. The data suggest statistically significant subjective improvement versus placebo at the 2 mg dosage for all parameters assessed.

Discussion

Cohn[152] compared estazolam 1 mg and 2 mg to flurazepam and placebo in approximately 100 adults with chronic sleep onset and maintenance insomnia in a parallel group design. Subjects were randomized to receive drug or placebo for seven consecutive nights. Outcomes were measured by sleep questionnaires (interval ratings and Likert scales). Dominguez[153] evaluated a similar population of 45 adults with estazolam 2 mg, flurazepam or placebo for 7 nights. Sleep variables were assessed by patient questionnaire. Scharf[154] studied 243 outpatients with complaints of sleep onset or maintenance difficulty. Subjects were randomly assigned to one of three parallel groups: estazolam 2 mg, flurazepam 30 mg or placebo. Medications were administered for 7 nights. Subjects rated sleep latency, TST, QOS and NOA on numerical interval questionnaires.

Three studies found statistically significant improvement in SL on patient ratings with estazolam 2 mg. The only study which included estazolam 1 mg reported no significant improvement on SL. All three studies reported significant improvement versus placebo in sleep duration at 2 mg. The 1 mg dosage also produced significant improvement in sleep duration. Sleep quality was likewise improved at both dosages studied, as were NOA. No studies assessed WASO or SE.

Quazepam

Summary

Seven studies evaluated the efficacy of quazepam versus placebo in randomized, controlled trials.[108,155–160] One of these studies[160] reported PSG findings while the remainder relied exclusively on subjective data derived from sleep questionnaires. Data analysis varies somewhat across these studies, rendering comparisons difficult. Only one investigation[160] met requirements for meta-analysis. Overall, the studies suggest efficacy in reducing time to onset of sleep, increasing TST, and reducing NOA. The methodologies employed were not comparable to the standard of data reporting required by GRADE and, therefore, no specific recommendation was made. Quazepam and its metabolites have long half-lives, raising concerns regarding accumulation and daytime impairment. Data regarding daytime sleepiness from these studies suggests a higher percentage of patients with somnolence in the active treatment group versus placebo, particularly at the 30 mg dosage.

Discussion

Alden[155] evaluated 57 insomnia subjects in a 5 night, parallel group design with quazepam 30 mg as the active drug. This study and all additional quazepam studies reported here (with the exception of Roth[160]) utilized patient sleep questionnaire data consisting of numerical interval and other rating scales. Hernandez[156] studied 36 insomnia outpatients with quazepam 15 mg and placebo in a similar five night design. Martinez[157] assessed 41 older adults (> 65 years) with insomnia in a controlled trial with quazepam 15 mg or placebo administered over 5 consecutive nights. Mendels[158] assessed the same dosage in 60 adult insomnia outpatients for five nights. O'Hair[159] reported results of a five night trial in 60 subjects with quazepam 30 mg. Scharf[108] studied quazepam 15 mg and triazolam 0.5 mg versus placebo over a five week period. During this time, subjects received active drug or placebo for nine consecutive nights, followed by 14 nights of every other night administration. Subjects were 65 insomnia outpatients. Finally, Roth[160] evaluated quazepam 7.5 mg and 15 mg versus placebo in 30 older insomnia subjects (> 60 years). PSG was conducted for two nights in the early phase of treatment (nights 1 and 2 of active treatment) and during the late phase (nights 6 and 7).

SLEEP LATENCY: Utilizing a cutoff of sleep latency < 45 min to identify "responders," Aden[155] reported quazepam 30 mg to be statistically superior to placebo. O'Hair[159] demonstrated quazepam 30 mg to be significantly better than placebo on an interval scale for sleep latency.

Hernandez[156] found quazepam 15 mg significantly better than placebo on sleep latency interval scales. Likewise, Scharf[108] reported significantly shorter latencies at this dosage on interval scales during active treatment nights in every-other-night administration although this was apparently not the case during the initial nightly administration. Using a 45 min sleep latency cutoff as described above,[155] Martinez[157] demonstrated a significantly higher percentage of responders to 15 mg in a geriatric population. Roth[160] did not report significant differences between quazepam 7.5 mg or 15 mg and placebo on PSG SL.

TOTAL SLEEP TIME: Utilizing a cutoff of sleep duration > 6 h to identify "responders," Aden[155] reported quazepam 30 mg to be statistically superior to placebo. O'Hair[159] demonstrated quazepam 30 mg to be significantly better than placebo on an interval scale for TST.

Hernandez[156] found quazepam 15 mg to be significantly superior to placebo on sleep duration interval scales. Likewise, Scharf[108] reported significantly longer duration at this dosage on interval scales during active treatment nights in every-other-night administration, except on the initial night of administration. Using a > 6 h sleep duration cutoff as described above,[155] Martinez[157] demonstrated a significantly higher percentage of responders to 15 mg in a geriatric population. Roth[160] reported improvement in PSG TST during early (treatment nights 1 and 2) and late (nights 6 and 7) with quazepam 15 mg in a geriatric insomnia population. A statistically significant effect with quazepam 7.5 mg was seen only during nights 6 and 7.

WAKE AFTER SLEEP ONSET: No studies reported placebo comparisons for WASO.

QUALITY OF SLEEP: The majority of studies of "sleep quality" with quazepam utilized a composite index for sleep quality (including questions on nightmares and overall evaluation of the medication) which is not consistent with sleep quality measures used in other studies; therefore these results are not

discussed. Scharf[108] reported a single measure of sleep quality ("How would you describe your sleep"). Quazepam 15 mg was significantly better than placebo on active treatment nights in both the nightly and every other night administration.

SLEEP EFFICIENCY: No studies reported placebo comparison data for SE.

NUMBER OF AWAKENINGS: Employing a threshold for "response" of < 2 awakenings, Aden[155] reported a significantly higher percentage of responders to quazepam 30 mg than placebo. O'Hair[159] also found significantly fewer awakenings at this dosage compared to placebo using interval scales. At the 15 mg dosage, two studies[156,157] found a significantly greater number of "responders" (i.e. < 2 awakenings) compared to placebo. No PSG data for NOA were reported.

ADVERSE EFFECTS: Five studies reported specific data for daytime somnolence. Aden[155] found an approximately four-fold higher rate of somnolence at 30 mg (quazepam 16/24; placebo = 4/26). At the same dose, O'Hair[159] reported somnolence in 12/30 quazepam and 5/30 placebo subjects. At 15 mg, Martinez[157] noted no difference in adverse events. Hernandez[156] reported somnolence in 9/30 quazepam subjects and 6/30 placebo subjects. Mendels[158] found 7/30 quazepam subjects and 4/30 placebo subjects demonstrated daytime somnolence.

Flurazepam

Summary

Sixteen studies met general inclusion and exclusion criteria.[98,100,101,105,109,149,152–154,161–167] No studies contained data adequate for meta-analysis. No meta-analysis of harms was possible. These studies were highly varied in design. Of these, three[100,101,105] included no flurazepam/placebo comparison and were excluded from discussion. All of the studies included one or both of the standard flurazepam doses: 15 mg and 30 mg.

Studies of the efficacy of flurazepam had numerous methodological inconsistencies, including instruments for subjective assessments of sleep outcomes that were highly variable across these studies, which made valid comparisons across studies impossible. Many studies incorporated interval scales with no reports of specific values. In light of these inconsistencies, and the related unavailability of meta-analyses, no recommendations regarding efficacy of flurazepam were made. The data for sleep onset at both the 15 mg and 30 mg dosages are mixed. The majority of studies did report increases in TST with the 30 mg dosage, but not at 15 mg. Data for WASO are limited to two studies, one of which (a PSG study) showed improvement at 30 mg. Sleep quality reports uniformly indicated improvement at both dosages, while reports for NOA suggest reduction at the 30 mg dosage only.

Discussion

Cohn[152] compared flurazepam 30 mg and placebo in approximately 100 adults with chronic sleep onset and maintenance insomnia in a parallel group design. The study, with a total n = 223, also included two dosages of estazolam, discussed elsewhere. Subjects were randomized to receive drug or placebo for seven consecutive nights. Outcomes were measured by sleep questionnaires (interval ratings and Likert scales). Dominguez[153] evaluated a similar population of 45 adults with flurazepam 30 mg or placebo for 7 nights. Sleep variables were assessed by patient questionnaire. Elie[161] studied 60 outpatient insomnia patients using a cross-over study design in which each patient received a single dose of five different drugs (or drug dosages) or placebo on one night of the week over a five consecutive week period. Study drugs included flurazepam 15 mg, three crossover dosages of loprazolam, and placebo. Outcomes included an index for sleep-onset based on patient questionnaires. Elie[162] investigated efficacy of flurazepam 30 mg and zopiclone versus placebo over 4 weeks. Flurazepam and placebo groups included 12 chronic insomnia patients per group. Subjects reported sleep outcome variables on post-sleep numerical rating questionnaires. Hartmann[162] studied 96 adult patients (n = 45 for flurazepam and placebo groups) with various insomnia complaints. Subjects were randomly assigned to receive flurazepam 30 mg, secobarbital, l-tryptophan, or placebo for one week of active treatment. Outcomes were assess by sleep logs which included subjective estimates of SL, NOA, duration of awakenings, and QOS.

Mamelak[164] investigated the effects of flurazepam 30 mg and zopiclone versus placebo in three groups of 10 insomnia subjects per group, each of which received one of the three treatment conditions for 12 consecutive nights. Subjective estimates of SL, TST and NOA were reported. Mamelak[165] studied 36 elderly patients with chronic insomnia. Patients were randomized to flurazepam 15 mg, brotizolam or placebo for 14 nights. Outcomes included patient-reported SL, NOA, TST and wake time. Daytime performance measures were conducted at the beginning of treatment and following conclusion. Melo de Paula[166] evaluated flurazepam 30 mg versus placebo and two dosages of lormetazepam in 60 adults with sleep onset or maintenance problems. Subjects received one of the four treatment conditions for two weeks. Outcome data included subjective SL, NOA and TST.

Reeves[98] investigated the efficacy of flurazepam 15 mg and triazolam versus placebo in 61 geriatric subjects (n = 27 for flurazepam and placebo groups) with sleep onset or maintenance insomnia. Subjective sleep outcomes were assessed by interval rating questionnaires. Salkind[167] evaluated flurazepam 15 and 30 mg versus placebo in 30 general practice insomnia patients. Subjects received each dose of flurazepam and placebo for one week in a crossover trial. Patient-reported SL, TST and QOS were primary outcome variables. Daytime residual effects were also reported. Scharf[154] studied 243 outpatients (n = 163 for flurazepam versus placebo) with complaints of sleep onset or maintenance difficulty. Subjects were randomly assigned to one of three parallel groups: flurazepam 30 mg, estazolam 2 mg or placebo. Treatments were administered for 7 nights. Subjects rated sleep latency, TST, QOS and NOA on numerical interval questionnaires. Sunshine[109] investigated the effects of 15 mg and 30 mg flurazepam versus two dosages of triazolam and placebo in a five-night crossover study, with subjects receiving each intervention for one night. Subjects were 25 inpatients who

complained of sleep onset and maintenance problems. Patients completed sleep questionnaires with interval ratings for TST and NOA.

Kripke[163] conducted the only identified PSG study of flurazepam. In this study, 99 subjects with chronic insomnia were randomized to one of four parallel groups (flurazepam 15 mg, flurazepam 30 mg, midazolam or placebo). Subjects received treatment for 14 consecutive nights, with PSG recordings on nights 1, 2, 7, 13 and 14. Objective SL, WASO, TST, and SE were reported.

SLEEP LATENCY: The only PSG study of 30 mg[163] found no significant reduction in SL versus placebo.

Five studies[152,154,164,166,167] reported statistically significant improvement on subjective ratings of sleep onset for flurazepam 30 mg versus placebo. Kripke[163] found improvement in patient-reported SL for 30 mg only in the early period (nights 1 and 2) of administration. No significant difference from placebo was evident at end of 14-day treatment. Four reports[109,149,153,162] found no significant subjective improvement in sleep onset with flurazepam 30 mg versus placebo.

Three studies[98,161,167] reported subjectively improved onset at the 15 mg dosage. Kripke[163] found patient-reported improvement at this dosage only on nights 1 and 2. Two investigations demonstrated no improvement in sleep onset for flurazepam 15 mg versus placebo.

TOTAL SLEEP TIME: Eight studies[109,152–154,163,164,166,167] reported statistically significant improvement for flurazepam 30 mg versus placebo on various subjective scales for sleep duration. One study[162] reported no significant improvement in duration at this dosage.

Two studies[109,167] found significantly improved patient-reported duration at the 15 mg dosage; Kripke[163] reported subjective improvement only on nights 1 and 2. Likewise, two studies[98,165] found no significant subjective improvement in sleep duration for flurazepam 15 mg.

WAKE AFTER SLEEP ONSET: Two studies reported data for WASO. Kripke[163] found significantly reduced PSG WASO with flurazepam 30 mg versus placebo. Mamelak[165] reported no significant reduction in subjective WASO with flurazepam 15 mg in an elderly insomnia population.

QUALITY OF SLEEP: Utilizing a variety of self-report scales, six studies[98,152–154,161,167] reported improvement in sleep quality with flurazepam versus placebo. Four studies[152–154,167] found improvement at the 30 mg dosage and three studies[98,161,167] at the 15 mg level.

SLEEP EFFICIENCY: One study[163] reported PSG sleep efficiency. Flurazepam 30 mg significantly improved sleep efficiency versus placebo.

NUMBER OF AWAKENINGS: Six studies[109,152–154,162,164] assessed subjective NOA with flurazepam 30 mg. All found significant reduction in NOA. Three studies[98,109,165] found no significant reduction in NOA with flurazepam 15 mg.

ADVERSE EFFECTS: Cohn[152] reported that 68% of flurazepam 30 mg subjects experienced an adverse event versus 43% of subjects receiving placebo. Approximately 50% of the flurazepam group reported somnolence, about twice the rate in the placebo population. Dominguez[153] found a significant increase in side effects for flurazepam 30 mg compared to placebo and stated that 73% of side effects described as "undetermined" were reports of somnolence. Elie[161] indicated that there was no significant difference in adverse events between flurazepam 15 mg and placebo; likewise Elie[162] found no difference in rates of somnolence for flurazepam 30 mg versus placebo. Mamelak[164] found significant performance impairment with flurazepam 30 mg. Mamelak[165] reported significantly shorter latencies to sleep on MSLT at the beginning and end of treatment. The authors also found significant impairment on digit symbol substitution and serial learning as well as a significantly slower rate of improvement on reaction, response and movement time. Divided attention was also impaired at end of treatment. Reeves[98] noted that 6 of 13 flurazepam subjects reported somnolence (versus 4/14 in the placebo group). Salkind[167] described impaired motor performance in the flurazepam 30 mg group (although not in the 15 mg group) and a significantly higher rate of "hangover effect" at the higher dosage. In the cross-over design, 11 of 30 flurazepam group experienced morning drowsiness/hangover, which was reported by only 3 of 30 subjects during the flurazepam 15 mg period and 2 of 30 while taking placebo. Finally, Scharf[154] found AEs in 73% of the flurazepam 30 mg group versus 43% on placebo subjects. Somnolence was the most common event, reported by 57% of flurazepam subjects and 23% of the placebo group.

Oxazepam

Götestam[168] studied the efficacy of oxazepam 25 mg vs. placebo with a crossover design in 28 patients with "insomnia." Subjective reports using interval ratings showed a significant reduction in SL and significant improvement in QOS.

Quetiapine

One study[169] investigated the efficacy of quetiapine versus placebo control in primary insomnia. However, the study included only 13 subjects. Numerical increase in subjective TST and decrease in subjective SL were found, but these differences were not statistically significant, possibly due to small sample size.

Gabapentin

One study[170] evaluated gabapentin for treatment of primary insomnia. This was an open-label investigation with 18 subjects, variable dosages, and no placebo control. Therefore, the trial was excluded.

Paroxetine

Two studies assessed paroxetine for treatment of primary insomnia. Nowell[171] reported a trial of variable dosage in 15 patients, without placebo control. As a result, this investigation was excluded.

Reynolds[172] evaluated paroxetine 10 mg/20 mg in 27 older adults with primary insomnia who were randomized to drug or placebo. The two doses were pooled for statistical analysis.

PSG data showed a modest but significant *increase* in SL, decrease in WASO and no difference in SE versus placebo. Sleep quality was improved.

Trimipramine

Hohagen[173] studied the effects of trimipramine in 15 adults with primary insomnia. No placebo control was included and, as a result, the study was excluded. Riemann[174] evaluated 55 adults with primary insomnia in a placebo-controlled double blind study. Dosage was variable (50–200 mg; mean 109.4 mg), but pooled for analysis. No significant difference was observed between trimipramine and placebo for PSG TST or SL, but SE was significantly improved with trimipramine. Subjective sleep quality also showed significant improvement.

DISCUSSION AND FUTURE DIRECTIONS

Defining "Efficacy"

Assessment of the efficacy of a given agent for the treatment of chronic insomnia is a complex and challenging task. It remains unclear which are the most important variables for defining efficacy. Older studies, particularly the majority of investigations of benzodiazepine efficacy, utilized a variety of predominantly subjective scales and questionnaires. These are highly diverse and did not often include specific numerical patient estimates for sleep outcomes. Since the advent of newer benzodiazepine receptor agonists, more specific and uniform outcomes for both patient-reported and objective outcomes (e.g., self-reported and PSG sleep onset latency, wake time after sleep onset, and total sleep time) have been employed, although continued substantial variability in data reporting has not been uncommon.

In addition to the variability in outcome measures reported, there are a number of critical unresolved issues regarding evaluating the efficacy of treatments for chronic insomnia. One is the relative importance of subjective versus objective data. Another is whether metrics of sleep quality, whether they be subjective or objective (e.g. analysis of the microstructure of sleep or related physiological parameters), are perhaps more pertinent than measures of SL, TST or WASO. An additional issue of importance is whether efficacy is better reflected by measures of daytime alertness and cognitive, emotional, and psychomotor function than by measures of sleep. Recent behavioral treatment studies in chronic insomnia have taken yet another direction: measuring response or remission of the insomnia syndrome as the most clinically-relevant outcome. This approach makes sense from a patient-centered approach, since most patients complain of "difficulty" falling asleep or staying asleep, rather than tying their complaints to any specific numerical value. Indeed, several studies have identified a group of "non-complaining poor sleepers" whose quantitative sleep measures are similar to those with insomnia. Examining the insomnia syndrome is also useful because it addresses both sleep-related and wake-related symptoms.

Absent clear answers to these questions, the present analysis relies on conventional subjective and objective measures of major sleep variables (sleep onset latency, total sleep time or wake time after sleep onset). The meta-analyses conducted yield recommendations for use of a limited number of drugs for a limited number of specific indications (i.e. sleep onset and/or sleep maintenance). In all cases, the recommendations are "weak," in that they are based on relatively limited and low quality evidence. The majority of medications included in these analyses are FDA-approved drugs for treating insomnia. This is not surprising, given that FDA approval rests on the demonstration of statistically significant changes in both subjective and objective outcomes. Furthermore, FDA approval is based on standards of significant improvement versus placebo for one or more indications (i.e. sleep onset or sleep maintenance insomnia). Many agents, including some which are not FDA-approved hypnotics, have been shown in one or more studies to be "statistically significantly superior" to placebo for a given outcome(s), but are nonetheless *not* recommended for treatment of chronic insomnia in this guideline. It is important to understand the discrepancy between (1) FDA approval and/or demonstration of "statistically significant superiority" to placebo and (2) the recommendations included in this publication. The discrepancy results from different criteria employed by the FDA and individual studies, on the one hand, and the GRADE approach to clinical guidelines, on the other. The GRADE approach establishes evidence quality ratings and clinical significance thresholds that are not employed in individual research studies and FDA assessment for approval. The thresholds were determined by clinical judgement of the task force and represent best estimates of the degree of improvement which the "typical patient" would find significant. Although these thresholds are consistent with numerical values that have been recommended as thresholds in contemporary publications, these standards entail a certain amount of subjectivity on the part of the task force, as there are no data which suggest absolute standards for clinical significance. Without question, there may be divergent opinions regarding what constitutes clinical significance and efficacy. Indeed, the task force assumed that their recommendations are not absolute indications of the presence or absence of clinical utility of a given medication, but reflect their best judgment based on the available data. Each prescriber bears the responsibility for making treatment determinations with this in mind.

Patient selection and inclusion criteria for studies are variable and may substantially impact results for a given outcome (e.g. see Krystal, 2012). Studies not requiring a minimum inclusion criterion for a specific outcome (e.g. inclusion thresholds for SL or WASO) may be underpowered to identify significant change for that outcome. On the other hand, studies with stringent PSG criteria for inclusion may not represent the larger population of insomnia patients.

Understanding the Methodology

The recommendations of the task force were developed with the use of GRADE, a state-of-the-art methodology for assessment of clinical data. This approach has distinct strengths, as well as certain limitations. GRADE is a rigorous, detailed, and transparent system for evaluation of the relative strengths of evidence for a given intervention. It incorporates several considerations which may impact the quality of evidence for a treatment approach. These factors include the heterogeneity of

data (i.e. the degree of inconsistency of results across studies), imprecision of the data (i.e. 95% CI which cross the clinical significance threshold) and potential publication bias (as a result of industry sponsorship). Quality of evidence grades for randomized clinical trials begin at HIGH and are downgraded progressively for heterogeneity, imprecision, and/or potential publication bias. Since the vast majority of studies in this field are industry sponsored, the quality of evidence for nearly all of these studies is, therefore, reduced from HIGH to MODERATE. This is to be expected for clinical trials for many drugs (i.e. not only hypnotics), since the vast majority are industry-sponsored FDA registration studies. The extent to which this downgrading of evidence is warranted due to actual publication bias is unknown, but under the GRADE system we have chosen to adopt the conservative approach and assume risk of bias. When heterogeneity and imprecision are accounted for, the quality of evidence for many treatments considered is LOW or VERY LOW. These latter two factors are not uncommon, as there is substantial variability in sleep outcome variables across studies and confidence intervals frequently overlap the clinical thresholds for significance.

Meta-analysis requires specific data (numerical data for a given outcome, presented as mean and standard deviation). Many studies, particularly older investigations, do not report data in the required format. Some newer publications do not report data in this format because some sleep variables, particularly sleep onset latency, are not normally distributed. In this case, the preferred measure of central tendency is not the mean but the median, the standard deviation may not be a valid measure of the degree of dispersion, and the statistical analyses carried out are not based on the mean and standard deviation. The result of this is exclusion of substantial amounts of data from the formal meta-analyses. While these studies are discussed in the paper and (secondarily) considered in formulation of recommendations, the inability to include such data in meta-analysis represents a distinct limitation.

As described in the methodology section, GRADE requires a recommendation "for" or "against" use of each treatment. When the evidence for efficacy is clear-cut, with (1) relatively high quality of evidence; (2) a high degree of confidence that benefits clearly outweigh harms; and (3) evidence that the effects of treatment are of substantial magnitude, without imposition of significant burden to the patient, a "strong" recommendation is delivered in the form of, "we *recommend* clinicians use X for the treatment of chronic insomnia." When evidence for benefit is less clear and the quality lower, a "weak" recommendation is made in the form of, "We *suggest* that clinicians use (or not use)…" However, it is important for clinicians to understand that a recommendation against use, particularly when associated with low quality evidence, **is not equivalent to a demonstration of ineffectiveness**. Rather, it is often an indication that the available evidence is simply insufficient and fails to provide convincing support in favor of usage by GRADE standards. In the case of drugs (most commonly older drugs) for which none of the data were reported in a format amenable to meta-analysis, we refrain from making any recommendation. The specific indications for use of a hypnotic employed in this report are limited to "sleep onset" and "sleep maintenance." insomnia. We chose these since, from a practical clinical consideration, these are the primary complaints with which chronic insomnia patients present, and for which clinicians prescribe medication. Moreover, these are the subtypes of insomnia that were actually studied in many investigations, consistent with FDA approval strategies and the matching of drugs to particular types of sleep disturbance.

Hence, some medications may show substantial improvement in TST or sleep quality, yet demonstrate no or insignificant reduction in SL, WASO or NOA to qualify for a recommendation in favor of use.

As described, we established thresholds for clinically significant improvement for each objective and subjective major sleep outcome. Nevertheless, some degree of judgment was introduced in formulating final recommendations. For example, a medication may not have exceeded significance thresholds for both subjective and objective evidence but, when the totality of evidence (including those investigations which could not be included in the meta-analysis) was considered, the task force concluded that a reasonable standard had been met. These considerations also include the role of adverse effects in the decisions made.

Beyond the quality of evidence for or against use of a given drug for sleep onset or maintenance insomnia, the task force also considered the relative benefit:harm ratio and the likelihood that an informed patient would use a specific agent. To a great extent, these decisions are based on clinical judgement. With respect to the benefit:harm consideration, the data on adverse events is often limited or non-existent. This may reflect the fact that treatment-emergent adverse events (TEAEs) are typically not collected using specific assessment forms, but rather, rely on spontaneous reporting by research participants. In addition, the frequency of some TEAEs is so low that the reported studies are underpowered to find a difference from placebo. This also implies that the effect size for a TEAE would be very small, and hence, it is unlikely that the clinical significance of TEAEs has been underestimated. However, some TEAEs are very infrequent but very serious when they do occur (e.g. sleep-related behaviors with BzRA). Clinical trials are likely to underestimate such risks due to the limited number of patients treated and the limited duration of treatment. As a result of these considerations, assessment of potential harms is largely derived from clinical experience and theoretical considerations, rather than well-documented evidence. This is clearly a limitation of the analysis and further, more systematic investigation of adverse effects is necessary.

Prior to formulation of the specific recommendations, the task force—based on its clinical judgement and experience—indicated what medications well-informed patients would or would not choose to use. These judgments do not reflect the input of actual patients, but only the task force's judgment. In most cases, these judgments were in agreement with recommendations (i.e. an informed patient is likely to use a drug that is recommended and not likely to use one that is not). In certain cases (e.g. melatonin), the task force considered that, given widespread use and apparently benign side effect profiles, informed patients may be likely to use a specific drug even when data do not clearly support a recommendation for use.

Clinical Application

Administration of sleep-promoting medication for chronic insomnia is one possible component of what must be a comprehensive approach to evaluation and treatment of chronic insomnia. This approach must include adequate assessment of cause and characteristics of the disorder as well as evaluation and treatment of contributing comorbidities. The latter may include any one or more of numerous medical, neurological and mental disorders, as well as other primary sleep disorders.

Numerous investigations have demonstrated that hypnotic medications are comparably efficacious to CBT-I during acute treatment.[11,112,175] However, these studies also make clear that the gains associated with CBT-I are durable following completion of treatment, whereas those associated with medication tend to dissipate following discontinuation of the drug. The vast majority of investigations which are included in the current analysis address relatively short-term use (e.g. one day to five weeks). Some studies have shown that long-term treatment with at least newer generation BzRA hypnotics can be safe and effective under properly controlled conditions. However, chronic use should be reserved for those individuals for whom CBT is either inaccessible or ineffective, who have been appropriately screened for contraindications to such treatment, who maintain long-term gains with medication, and who are followed regularly. Patient preference must also be considered in the determination of treatment approach.

The investigations which are included in this analysis were focused on "primary" chronic insomnia, with the exception of some older studies (e.g. zaleplon) which included some patients with "mild" mental disorders, The extent to which these findings apply to chronic insomnia associated with major comorbidities is uncertain, although a limited number of comparative studies suggest at least some degree of efficacy in such cases. It should also be emphasized that the findings presented in this report apply only to adults. None of the agents discussed in this report are approved for use in children and none of the findings presented apply to children or adolescents. There is very little information concerning pharmacotherapy for childhood insomnia. Although independent analyses of efficacy in older adults were not conducted, examination of the findings suggests comparable efficacy across the adult age range. Pharmacokinetic and pharmacodynamic properties of many medications, including benzodiazepine receptor agonist drugs, differ among older and younger adults, necessitating lower starting dosages. The limited information from these studies regarding adverse effects in older adults does not allow meaningful conclusions about the frequency of such events in older patients compared to a younger population. The American Geriatric Society Beers criteria recommend that benzodiazepines be avoided for treatment of insomnia in older patients, due to risk of cognitive impairment, falls, and motor vehicle accidents. The criteria further recommend that newer generation benzodiazepine receptor agonists be limited to shorter-term use (< 90 days).

The data on adverse effects derived from these clinical trials, in general, do not suggest a high frequency of serious side effects. However, the data are scant and inconsistent, suggesting that caution should be applied in the assessment of relative risks associated with use of hypnotic medications. Other reported adverse effects include—but are not limited to—dependency/withdrawal, cognitive impairment, falls/fractures, parasomnias, and driving impairment and motor vehicle accidents. Epidemiological studies have also suggested a possible link between hypnotic use and infection, depression and overall mortality risk. These complications are observed most frequently in older populations, who are among the most frequent users of these drugs. Risks of dependency and serious withdrawal complications are of greatest concern with true benzodiazepine agents, particularly in the setting of escalating, long-term usage and insufficient monitoring. However, although much concern has understandably been raised about potential tolerance and addiction to these drugs, there is limited information regarding the true incidence of these complications. The risks associated with use of these agents are clearly increased not only in the elderly but also when they are used in dosages in excess of those recommended, or when combined with other psychoactive agents.[38] Given the known sedative effects of these agents, particularly those with longer half-lives, clinicians must be diligent in cautioning patients regarding potential complications related to sedation. Such complications are most likely to occur with longer-acting agents and during morning hours following bedtime administration. Use of shorter-acting agents and the lowest effective dosage may help to reduce sedation-related complications. Appropriate patient counseling and careful monitoring will also serve to minimize risk. Complete avoidance of these medications should also be considered in those who may be particularly susceptible to adverse outcomes.

Future Directions

In an attempt to develop meaningful clinical practice recommendations for the use of sleep-promoting medications, it became increasingly clear to the task force that this endeavor is fraught with multiple limitations. While existing data (especially more recent data) provide a reasonable foundation for certain recommendations contained in this study, the overall quality of evidence is relatively low in the vast majority of cases. For numerous drugs, there is simply insufficient evidence available to draw on in determining whether or not a compound is efficacious. Data reporting, especially that of older studies, is highly variable and idiosyncratic. As a result, comparing data from one study to another, or conducting meta-analyses of data, is not possible. Virtually all studies of prescription hypnotic agents are industry-funded. While the reasons for this are understandable, the potential for publication bias, particularly lack of publication of negative results, compromises the quality of evidence to a significant degree. Moreover, the role of industry in study design and data analysis may further compromise uniformity of data reporting.

With these limitations in mind, the task force recommends the following for future investigations:

1. Clear definitions of inclusion and exclusion criteria;
2. Adequately powering studies to detect significant differences for key sleep variables;
3. Development and utilization of uniform data collection instruments which will promote improved cross-study analysis and comparisons;

4. Standardized statistical analysis and data presentation. The majority of newer investigations now present means ± SD for specific PSG or sleep diary data. For those variables that are not normally distributed, a transformation can be sought which converts the probability distribution to the normal distribution and the transformed mean and SD can then be reported. An effort to report means and SD data should be made for all studies;
5. Although specific numerical data for individual sleep are useful in assessing the efficacy of pharmacological treatment for insomnia, other approaches to such evaluation may be more clinically meaningful. Specifically, determination of the efficacy of a drug in achieving remission of chronic insomnia disorder has been employed in cognitive behavioral treatments for insomnia and should be considered as a clinically relevant outcome in pharmacological trials. This may include not only subjective and objective outcome data for major sleep outcomes, but also sleep quality and daytime functional outcomes;
6. To the extent possible, encourage funding for independent, non-industry investigation of the efficacy and effectiveness of hypnotic medications;
7. Data for adverse events associated with hypnotic medications are not collected and analyzed in standard ways. This is a widespread problem common to studies of all types of medications. Continued efforts should be made to standardize and systematize the reporting of adverse effects data;
8. Daytime sedation, with concomitant risk of motor vehicle or occupational accidents, is a significant potential risk. Further efforts to include objective assessments of performance impairments which may be associated with daytime sedation is encouraged;
9. Virtually no data exists regarding the use of sleep-promoting agents in children. Yet, such medications are not infrequently used in this age group. As such, studies of the efficacy and safety of sleep-promoting medications in children and adolescents should be required.

Summary

This analysis is, to the best of our knowledge, the most comprehensive assessment of efficacy of individual sleep-promoting agents published to date. It relies heavily on rigorous evaluation of the quality of evidence for efficacy, based on GRADE, as well as determination of potential adverse effects, to the extent possible. It is intended to serve as a useful guide for clinicians in prescribing medications for the treatment of chronic insomnia. This analysis, however, also makes it abundantly clear that the availability and quality of the data which serve as the foundation for such recommendations are sorely limited. The result is that many commonly used drugs, including some which carry FDA approval for treatment of insomnia, are not recommended. Further data are required to formulate any reasonable conclusion regarding their efficacy or lack thereof. As a result, clinicians must continue to exercise sound clinical judgment, based not only on these recommendations, but also on clinical experience, prior patient response, patient preferences, and potential adverse effects.

REFERENCES

1. 2005 Adult sleep habits and styles. National Sleep Foundation Web site. https://sleepfoundation.org/sites/default/files/2005_summary_of_findings.pdf. Published 2005. Accessed June 2016.
2. Kassam A, Patten SB. Hypnotic use in a population-based sample of over thirty-five thousand interviewed Canadians. *Popul Health Metr.* 2006;4:15.
3. Kaufmann CN, Spira AP, Alexander GC, Rutkow L, Mojtabai R. Trends in prescribing of sedative-hypnotic medications in the USA: 1993-2010. *Pharmacoepidemiol Drug Saf.* 2016;25(6):637–645.
4. Ohayon MM, Caulet M, Guilleminault C. How a general population perceives its sleep and how this relates to the complaint of insomnia. *Sleep.* 1997;20(9):715–723.
5. Chesson A, Jr., Hartse K, Anderson WM, et al. Practice parameters for the evaluation of chronic insomnia. An American Academy of Sleep Medicine report. Standards of Practice Committee of the American Academy of Sleep Medicine. *Sleep.* 2000;23(2):237–241.
6. Sateia MJ, Doghramji K, Hauri PJ, Morin CM. Evaluation of chronic insomnia. An American Academy of Sleep Medicine review. *Sleep.* 2000;23(2):243–308.
7. Schutte-Rodin S, Broch L, Buysse D, Dorsey C, Sateia M. Clinical guideline for the evaluation and management of chronic insomnia in adults. *J Clin Sleep Med.* 2008;4(5):487–504.
8. Chesson AL, Jr., Anderson WM, Littner M, et al. Practice parameters for the nonpharmacologic treatment of chronic insomnia. An American Academy of Sleep Medicine report. Standards of Practice Committee of the American Academy of Sleep Medicine. *Sleep.* 1999;22(8):1128–1133.
9. Morgenthaler T, Kramer M, Alessi C, et al. Practice parameters for the psychological and behavioral treatment of insomnia: an update. An American Academy of Sleep Medicine report. *Sleep.* 2006;29(11):1415–1419.
10. Morin CM, Bootzin RR, Buysse DJ, Edinger JD, Espie CA, Lichstein KL. Psychological and behavioral treatment of insomnia: update of the recent evidence (1998-2004). *Sleep.* 2006;29(11):1398–1414.
11. Morin CM, Colecchi C, Stone J, Sood R, Brink D. Behavioral and pharmacological therapies for late-life insomnia: a randomized controlled trial. *JAMA.* 1999;281(11):991–999.
12. Meolie AL, Rosen C, Kristo D, et al. Oral nonprescription treatment for insomnia: an evaluation of products with limited evidence. *J Clin Sleep Med.* 2005;1(2):173–187.
13. NIH State-of-the-Science Conference Statement on manifestations and management of chronic insomnia in adults. *NIH Consens State Sci Statements.* 2005;22(2):1–30.
14. Wilson SJ, Nutt DJ, Alford C, et al. British Association for Psychopharmacology consensus statement on evidence-based treatment of insomnia, parasomnias and circadian rhythm disorders. *J Psychopharmacol.* 2010;24(11):1577–1601.
15. Qaseem A, Kansagara D, Forciea MA, Cooke M, Denberg TD. Management of chronic insomnia disorder in adults: a clinical practice guideline from the American College of Physicians. *Ann Intern Med.* 2016;165(2):125–133.
16. American Academy of Sleep Medicine. *International Classification of Sleep Disorders.* 3rd ed. Darien, IL: American Academy of Sleep Medicine; 2014.
17. Ellis JG, Perlis ML, Neale LF, Espie CA, Bastien CH. The natural history of insomnia: focus on prevalence and incidence of acute insomnia. *J Psychiatr Res.* 2012;46(10):1278–1285.
18. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. *Sleep Med Rev.* 2002;6(2):97–111.
19. Ohayon MM. Observation of the natural evolution of insomnia in the american general population cohort. *Sleep Med Clin.* 2009;4(1):87–92.
20. DiBonaventura M, Richard L, Kumar M, Forsythe A, Flores NM, Moline M. The Association between Insomnia and Insomnia Treatment Side Effects on Health Status, Work Productivity, and Healthcare Resource Use. *PLoS One.* 2015;10(10):e0137117.

21. Spira AP, Kaufmann CN, Kasper JD, et al. Association between insomnia symptoms and functional status in U.S. older adults. *J Gerontol B Psychol Sci Soc Sci.* 2014;69 Suppl 1:S35–S41.
22. Leigh JP. Employee and job attributes as predictors of absenteeism in a national sample of workers: the importance of health and dangerous working conditions. *Soc Sci Med.* 1991;33(2):127–137.
23. Hagg SA, Toren K, Lindberg E. Role of sleep disturbances in occupational accidents among women. *Scand J Work Environ Health.* 2015;41(4):368–376.
24. Laugsand LE, Strand LB, Vatten LJ, Janszky I, Bjorngaard JH. Insomnia symptoms and risk for unintentional fatal injuries--the HUNT Study. *Sleep.* 2014;37(11):1777–1786.
25. Baglioni C, Battagliese G, Feige B, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *J Affect Disord.* 2011;135(1-3):10–19.
26. Breslau N, Roth T, Rosenthal L, Andreski P. Sleep disturbance and psychiatric disorders: a longitudinal epidemiological study of young adults. *Biol Psychiatry.* 1996;39(6):411–418.
27. Fernandez-Mendoza J, Vgontzas AN, Liao D, et al. Insomnia with objective short sleep duration and incident hypertension: the Penn State Cohort. *Hypertension.* 2012;60(4):929–935.
28. Bathgate CJ, Edinger JD, Wyatt JK, Krystal AD. Objective but not subjective short sleep duration associated with increased risk for hypertension in individuals with insomnia. *Sleep.* 2016;39(5):1037–1045.
29. Daley M, Morin CM, LeBlanc M, Gregoire JP, Savard J. The economic burden of insomnia: direct and indirect costs for individuals with insomnia syndrome, insomnia symptoms, and good sleepers. *Sleep.* 2009;32(1):55–64.
30. Ozminkowski RJ, Wang S, Walsh JK. The direct and indirect costs of untreated insomnia in adults in the United States. *Sleep.* 2007;30(3):263–273.
31. Sivertsen B, Lallukka T, Salo P. The economic burden of insomnia at the workplace. An opportunity and time for intervention? *Sleep.* 2011;34(9):1151–1152.
32. Leger D, Guilleminault C, Bader G, Levy E, Paillard M. Medical and socio-professional impact of insomnia. *Sleep.* 2002;25(6):625–629.
33. Kleinman NL, Brook RA, Doan JF, Melkonian AK, Baran RW. Health benefit costs and absenteeism due to insomnia from the employer's perspective: a retrospective, case-control, database study. *J Clin Psychiatry.* 2009;70(8):1098–1104.
34. Montgomery P, Dennis J. Cognitive behavioural interventions for sleep problems in adults aged 60+. *Cochrane Database Syst Rev.* 2002;(2):CD003161.
35. Trauer JM, Qian MY, Doyle JS, Rajaratnam SM, Cunnington D. Cognitive Behavioral Therapy for Chronic Insomnia: A Systematic Review and Meta-analysis. *Ann Intern Med.* 2015;163(3):191–204.
36. Wu JQ, Appleman ER, Salazar RD, Ong JC. Cognitive behavioral therapy for insomnia comorbid with psychiatric and medical conditions: a meta-analysis. *JAMA Intern Med.* 2015;175(9):1461–1472.
37. Zachariae R, Lyby MS, Ritterband LM, O'Toole MS. Efficacy of internet-delivered cognitive-behavioral therapy for insomnia - a systematic review and meta-analysis of randomized controlled trials. *Sleep Med Rev.* 2016;30:1–10.
38. Bertisch SM, Herzig SJ, Winkelman JW, Buettner C. National use of prescription medications for insomnia: NHANES 1999-2010. *Sleep.* 2014;37(2):343–349.
39. Balter MB, Uhlenhuth EH. The beneficial and adverse effects of hypnotics. *J Clin Psychiatry.* 1991;52 Suppl:16–23.
40. Roehrs T, Roth T. 'Hypnotic' prescription patterns in a large managed-care population. *Sleep Med.* 2004;5(5):463–466.
41. 2005 Adult Sleep Habits and Styles. National Sleep Foundation Web site. https://sleepfoundation.org/sleep-polls-data/sleep-in-america-poll/2005-adult-sleep-habits-and-styles. Published 2005. Accessed March 2016.
42. Cheung JM, Atternas K, Melchior M, Marshall NS, Fois RA, Saini B. Primary health care practitioner perspectives on the management of insomnia: a pilot study. *Aust J Prim Health.* 2014;20(1):103–112.
43. Pawaskar MD, Joish VN, Camacho FT, Rasu RS, Balkrishnan R. The influence of co-morbidities on prescribing pharmacotherapy for insomnia: evidence from US national outpatient data 1995-2004. *J Med Econ.* 2008;11(1):41–56.
44. Siriwardena AN, Apekey T, Tilling M, Dyas JV, Middleton H, Orner R. General practitioners' preferences for managing insomnia and opportunities for reducing hypnotic prescribing. *J Eval Clin Pract.* 2010;16(4):731–737.
45. Nowell PD, Mazumdar S, Buysse DJ, Dew MA, Reynolds CF, 3rd, Kupfer DJ. Benzodiazepines and zolpidem for chronic insomnia: a meta-analysis of treatment efficacy. *JAMA.* 1997;278(24):2170–2177.
46. Holbrook AM, Crowther R, Lotter A, Cheng C, King D. Meta-analysis of benzodiazepine use in the treatment of insomnia. *CMAJ.* 2000;162(2):225–233.
47. Dundar Y, Dodd S, Strobl J, Boland A, Dickson R, Walley T. Comparative efficacy of newer hypnotic drugs for the short-term management of insomnia: a systematic review and meta-analysis. *Hum Psychopharmacol.* 2004;19(5):305–322.
48. Glass J, Lanctot KL, Herrmann N, Sproule BA, Busto UE. Sedative hypnotics in older people with insomnia: meta-analysis of risks and benefits. *BMJ.* 2005;331(7526):1169.
49. Buscemi N, Vandermeer B, Friesen C, et al. The efficacy and safety of drug treatments for chronic insomnia in adults: a meta-analysis of RCTs. *J Gen Intern Med.* 2007;22(9):1335–1350.
50. Huedo-Medina TB, Kirsch I, Middlemass J, Klonizakis M, Siriwardena AN. Effectiveness of non-benzodiazepine hypnotics in treatment of adult insomnia: meta-analysis of data submitted to the Food and Drug Administration. *BMJ.* 2012;345:e8343.
51. Winkler A, Auer C, Doering BK, Rief W. Drug treatment of primary insomnia: a meta-analysis of polysomnographic randomized controlled trials. *CNS Drugs.* 2014;28(9):799–816.
52. Wilt TJ, MacDonald R, Brasure M, et al. Pharmacologic treatment of insomnia disorder: an evidence report for a clinical practice guideline by the American College of Physicians. *Ann Intern Med.* 2016;165(2):103–112.
53. Morgenthaler TI, Deriy L, Heald JL, Thomas SM. The evolution of the AASM clinical practice guidelines: another step forward. *J Clin Sleep Med.* 2016;12(1):129–135.
54. Herring WJ, Connor KM, Ivgy-May N, et al. Suvorexant in patients with insomnia: results from two 3-month randomized controlled clinical trials. *Biol Psychiatry.* 2016;79(2):136–148.
55. Herring WJ, Snyder E, Budd K, et al. Orexin receptor antagonism for treatment of insomnia: a randomized clinical trial of suvorexant. *Neurology.* 2012;79(23):2265–2274.
56. Ancoli-Israel S, Krystal AD, McCall WV, et al. A 12-week, randomized, double-blind, placebo-controlled study evaluating the effect of eszopiclone 2 mg on sleep/wake function in older adults with primary and comorbid insomnia. *Sleep.* 2010;33(2):225–234.
57. Erman MK, Zammit G, Rubens R, et al. A polysomnographic placebo-controlled evaluation of the efficacy and safety of eszopiclone relative to placebo and zolpidem in the treatment of primary insomnia. *J Clin Sleep Med.* 2008;4(3):229–234.
58. McCall WV, Erman M, Krystal AD, et al. A polysomnography study of eszopiclone in elderly patients with insomnia. *Curr Med Res Opin.* 2006;22(9):1633–1642.
59. Scharf M, Erman M, Rosenberg R, et al. A 2-week efficacy and safety study of eszopiclone in elderly patients with primary insomnia. *Sleep.* 2005;28(6):720–727.
60. Uchimura N, Kamijo A, Kuwahara H, et al. A randomized placebo-controlled polysomnographic study of eszopiclone in Japanese patients with primary insomnia. *Sleep Med.* 2012;13(10):1247–1253.
61. Zammit GK, McNabb LJ, Caron J, Amato DA, Roth T. Efficacy and safety of eszopiclone across 6-weeks of treatment for primary insomnia. *Curr Med Res Opin.* 2004;20(12):1979–1991.
62. Boyle J, Trick L, Johnsen S, Roach J, Rubens R. Next-day cognition, psychomotor function, and driving-related skills following nighttime administration of eszopiclone. *Hum Psychopharmacol.* 2008;23(5):385–397.
63. Krystal AD, Walsh JK, Laska E, et al. Sustained efficacy of eszopiclone over 6 months of nightly treatment: results of a randomized, double-blind, placebo-controlled study in adults with chronic insomnia. *Sleep.* 2003;26(7):793–799.
64. Walsh JK, Krystal AD, Amato DA, et al. Nightly treatment of primary insomnia with eszopiclone for six months: effect on sleep, quality of life, and work limitations. *Sleep.* 2007;30(8):959–968.

65. Soares CN, Joffe H, Rubens R, Caron J, Roth T, Cohen L. Eszopiclone in patients with insomnia during perimenopause and early postmenopause: a randomized controlled trial. *Obstet Gynecol.* 2006;108(6):1402–1410.
66. Joffe H, Petrillo L, Viguera A, et al. Eszopiclone improves insomnia and depressive and anxious symptoms in perimenopausal and postmenopausal women with hot flashes: a randomized, double-blinded, placebo-controlled crossover trial. *Am J Obstet Gynecol.* 2010;202(2):171.e1-e171.e111.
67. Hedner J, Yaeche R, Emilien G, Farr I, Salinas E. Zaleplon shortens subjective sleep latency and improves subjective sleep quality in elderly patients with insomnia. The Zaleplon Clinical Investigator Study Group. *Int J Geriatr Psychiatry.* 2000;15(8):704–712.
68. Walsh JK, Vogel GW, Scharf M, et al. A five week, polysomnographic assessment of zaleplon 10 mg for the treatment of primary insomnia. *Sleep Med.* 2000;1(1):41–49.
69. Ancoli-Israel S, Walsh JK, Mangano RM, Fujimori M. Zaleplon, A Novel Nonbenzodiazepine Hypnotic, Effectively Treats Insomnia in Elderly Patients Without Causing Rebound Effects. *Prim Care Companion J Clin Psychiatry.* 1999;1(4):114–120.
70. Elie R, Ruther E, Farr I, Emilien G, Salinas E. Sleep latency is shortened during 4 weeks of treatment with zaleplon, a novel nonbenzodiazepine hypnotic. Zaleplon Clinical Study Group. *J Clin Psychiatry.* 1999;60(8):536–544.
71. Fry J, Scharf M, Mangano R, Fujimori M. Zaleplon improves sleep without producing rebound effects in outpatients with insomnia. Zaleplon Clinical Study Group. *Int Clin Psychopharmacol.* 2000;15(3):141–152.
72. Dorsey CM, Lee KA, Scharf MB. Effect of zolpidem on sleep in women with perimenopausal and postmenopausal insomnia: a 4-week, randomized, multicenter, double-blind, placebo-controlled study. *Clin Ther.* 2004;26(10):1578–1586.
73. Herrmann WM, Kubicki ST, Boden S, Eich FX, Attali P, Coquelin JP. Pilot controlled double-blind study of the hypnotic effects of zolpidem in patients with chronic 'learned' insomnia: psychometric and polysomnographic evaluation. *J Int Med Res.* 1993;21(6):306–322.
74. Jacobs GD, Pace-Schott EF, Stickgold R, Otto MW. Cognitive behavior therapy and pharmacotherapy for insomnia: a randomized controlled trial and direct comparison. *Arch Intern Med.* 2004;164(17):1888–1896.
75. Perlis ML, McCall WV, Krystal AD, Walsh JK. Long-term, non-nightly administration of zolpidem in the treatment of patients with primary insomnia. *J Clin Psychiatry.* 2004;65(8):1128–1137.
76. Randall S, Roehrs TA, Roth T. Efficacy of eight months of nightly zolpidem: a prospective placebo-controlled study. *Sleep.* 2012;35(11):1551–1557.
77. Ware JC, Walsh JK, Scharf MB, Roehrs T, Roth T, Vogel GW. Minimal rebound insomnia after treatment with 10-mg zolpidem. *Clin Neuropharmacol.* 1997;20(2):116–125.
78. Walsh JK, Erman M, Erwin CW, et al. Subjective hypnotic efficacy of trazodone and zolpidem in DSMIII-R primary insomnia. *Hum Psychopharmacol.* 1998;13(3)191–198.
79. Staner L, Ertle S, Boeijinga P, et al. Next-day residual effects of hypnotics in DSM-IV primary insomnia: a driving simulator study with simultaneous electroencephalogram monitoring. *Psychopharmacology (Berl).* 2005;181(4):790–798.
80. Walsh JK, Soubrane C, Roth T. Efficacy and safety of zolpidem extended release in elderly primary insomnia patients. *Am J Geriatr Psychiatry.* 2008;16(1):44–57.
81. Roth T, Soubrane C, Titeux L, Walsh JK. Efficacy and safety of zolpidem-MR: a double-blind, placebo-controlled study in adults with primary insomnia. *Sleep Med.* 2006;7(5):397–406.
82. Scharf MB, Roth T, Vogel GW, Walsh JK. A multicenter, placebo-controlled study evaluating zolpidem in the treatment of chronic insomnia. *J Clin Psychiatry.* 1994;55(5):192–199.
83. Allain H, Arbus L, Schück S, Group ZS. Efficacy and safety of zolpidem administered 'as needed' in primary insomnia: results of a double-blind, placebo-controlled study. *Clin Drug Investig.* 2001;21(6):391–400.
84. Cluydts R, Peeters K, de Bouyalsky I, Lavoisy J. Comparison of continuous versus intermittent administration of zolpidem in chronic insomniacs: a double-blind, randomized pilot study. *J Int Med Res.* 1998;26(1):13–24.
85. Hajak G, Cluydts R, Declerck A, et al. Continuous versus non-nightly use of zolpidem in chronic insomnia: results of a large-scale, double-blind, randomized, outpatient study. *Int Clin Psychopharmacol.* 2002;17(1):9–17.
86. Krystal AD, Erman M, Zammit GK, Soubrane C, Roth T. Long-term efficacy and safety of zolpidem extended-release 12.5 mg, administered 3 to 7 nights per week for 24 weeks, in patients with chronic primary insomnia: a 6-month, randomized, double-blind, placebo-controlled, parallel-group, multicenter study. *Sleep.* 2008;31(1):79–90.
87. Maroo N, Hazra A, Das T. Efficacy and safety of a polyherbal sedative-hypnotic formulation NSF-3 in primary insomnia in comparison to zolpidem: a randomized controlled trial. *Indian J Pharmacol.* 2013;45(1):34–39.
88. Roth T, Hull SG, Lankford DA, Rosenberg R, Scharf MB. Low-dose sublingual zolpidem tartrate is associated with dose-related improvement in sleep onset and duration in insomnia characterized by middle-of-the-night (MOTN) awakenings. *Sleep.* 2008;31(9):1277–1284.
89. Roth T, Krystal A, Steinberg FJ, Singh NN, Moline M. Novel sublingual low-dose zolpidem tablet reduces latency to sleep onset following spontaneous middle-of-the-night awakening in insomnia in a randomized, double-blind, placebo-controlled, outpatient study. *Sleep.* 2013;36(2):189–196.
90. Staner C, Joly F, Jacquot N, et al. Sublingual zolpidem in early onset of sleep compared to oral zolpidem: polysomnographic study in patients with primary insomnia. *Curr Med Res Opin.* 2010;26(6):1423–1431.
91. Tu JH, Chung WC, Yang CY, Tzeng DS. A comparison between acupuncture versus zolpidem in the treatment of primary insomnia. *Asian J Psychiatr.* 2012;5(3):231–235.
92. Voshaar RC, van Balkom AJ, Zitman FG. Zolpidem is not superior to temazepam with respect to rebound insomnia: a controlled study. *Eur Neuropsychopharmacol.* 2004;14(4):301–306.
93. Walsh JK. Zolpidem "as needed" for the treatment of primary insomnia: a double-blind, placebo-controlled study. *Sleep Med Rev.* 2002;6 Suppl 1:S7–S10; discussion S10–S11, S31–S33.
94. Walsh JK, Roth T, Randazzo A, et al. Eight weeks of non-nightly use of zolpidem for primary insomnia. *Sleep.* 2000;23(8):1087–1096.
95. Zammit GK, Corser B, Doghramji K, et al. Sleep and residual sedation after administration of zaleplon, zolpidem, and placebo during experimental middle-of-the-night awakening. *J Clin Sleep Med.* 2006;2(4):417–423.
96. Roehrs T, Bonahoom A, Pedrosi B, Rosenthal L, Roth T. Treatment regimen and hypnotic self-administration. *Psychopharmacology (Berl).* 2001;155(1):11–17.
97. Elie R, Frenay M, Le Morvan P, Bourgouin J. Efficacy and safety of zopiclone and triazolam in the treatment of geriatric insomniacs. *Int Clin Psychopharmacol.* 1990;5 Suppl 2:39–46.
98. Reeves RL. Comparison of triazolam, flurazepam, and placebo as hypnotics in geriatric patients with insomnia. *J Clin Pharmacol.* 1977;17(5-6):319–323.
99. Autret E, Maillard F, Autret A. Comparison of the clinical hypnotic effects of zopiclone and triazolam. *Eur J Clin Pharmacol.* 1987;31(5):621–623.
100. Bowen AJ. Comparative efficacy of triazolam, flurazepam and placebo in out-patients insomniacs. *J Int Med Res.* 1978;6(4):337–342.
101. Fabre LF, Jr., Gross L, Pasigajen V, Metzler C. Multiclinic double-blind comparison of triazolam and flurazepam for seven nights in outpatients with insomnia. *J Clin Pharmacol.* 1977;17(7):402–409.
102. Fleming JA, McClure DJ, Mayes C, Phillips R, Bourgouin J. A comparison of the efficacy, safety and withdrawal effects of zopiclone and triazolam in the treatment of insomnia. *Int Clin Psychopharmacol.* 1990;5 Suppl 2:29–37.
103. Greenblatt DJ, Harmatz JS, Zinny MA, Shader RI. Effect of gradual withdrawal on the rebound sleep disorder after discontinuation of triazolam. *N Engl J Med.* 1987;317(12):722–728.
104. Hajak G, Clarenbach P, Fischer W, Haase W, Ruther E. Zopiclone improves sleep quality and daytime well-being in insomniac patients: comparison with triazolam, flunitrazepam and placebo. *Int Clin Psychopharmacol.* 1994;9(4):251–261.
105. Leibowitz M, Sunshine A. Long-term hypnotic efficacy and safety of triazolam and flurazepam. *J Clin Pharmacol.* 1978;18(5-6):302–309.
106. Monti JM, Attali P, Monti D, Zipfel A, de la Giclais B, Morselli PL. Zolpidem and rebound insomnia--a double-blind, controlled polysomnographic study in chronic insomniac patients. *Pharmacopsychiatry.* 1994;27(4):166–175.

107. Rickels K, Gingrich RL, Jr., Morris RJ, et al. Triazolam in insomniac family practice patients. *Clin Pharmacol Ther.* 1975;18(3):315–324.
108. Scharf MB. Feasibility of an every-other-night regimen in insomniac patients: subjective hypnotic effectiveness of quazepam, triazolam, and placebo. *J Clin Psychiatry.* 1993;54(1):33–38.
109. Sunshine A. Comparison of the hypnotic activity of triazolam, flurazepam hydrochloride, and placebo. *Clin Pharmacol Ther.* 1975;17(5):573–577.
110. Glass JR, Sproule BA, Herrmann N, Busto UE. Effects of 2-week treatment with temazepam and diphenhydramine in elderly insomniacs: a randomized, placebo-controlled trial. *J Clin Psychopharmacol.* 2008;28(2):182–188.
111. Hindmarch I. Effects of hypnotic and sleep-inducing drugs on objective assessments of human psychomotor performance and subjective appraisals of sleep and early morning behaviour. *Br J Clin Pharmacol.* 1979;8(1):43S–46S.
112. Wu R, Bao J, Zhang C, Deng J, Long C. Comparison of sleep condition and sleep-related psychological activity after cognitive-behavior and pharmacological therapy for chronic insomnia. *Psychother Psychosom.* 2006;75(4):220–228.
113. Cuanang JR, Limos L. Treatment of insomnia with temazepam: double-blind, placebo-controlled evaluation. *Clin Ther.* 1982;4(5):402–412.
114. Fillingim JM. Double-blind evaluation of the efficacy and safety of temazepam in outpatients with insomnia. *Br J Clin Pharmacol.* 1979;8(1):73S–77S.
115. Heffron WA, Roth P. Double-blind evaluation of the safety and hypnotic efficacy of temazepam in insomniac outpatients. *Br J Clin Pharmacol.* 1979;8(1):69S–72S.
116. Tuk B, Oberye JJ, Pieters MS, et al. Pharmacodynamics of temazepam in primary insomnia: assessment of the value of quantitative electroencephalography and saccadic eye movements in predicting improvement of sleep. *Clin Pharmacol Ther.* 1997;62(4):444–452.
117. Wilson SJ, Rich AS, Rich NC, Potokar J, Nutt DJ. Evaluation of actigraphy and automated telephoned questionnaires to assess hypnotic effects in insomnia. *Int Clin Psychopharmacol.* 2004;19(2):77–84.
118. Kohsaka M, Kanemura T, Taniguchi M, et al. Efficacy and tolerability of ramelteon in a double-blind, placebo-controlled, crossover study in Japanese patients with chronic primary insomnia. *Expert Rev Neurother.* 2011;11(10):1389–1397.
119. Mayer G, Wang-Weigand S, Roth-Schechter B, Lehmann R, Staner C, Partinen M. Efficacy and safety of 6-month nightly ramelteon administration in adults with chronic primary insomnia. *Sleep.* 2009;32(3):351–360.
120. Roth T, Seiden D, Wang-Weigand S, Zhang J. A 2-night, 3-period, crossover study of ramelteon's efficacy and safety in older adults with chronic insomnia. *Curr Med Res Opin.* 2007;23(5):1005–1014.
121. Zammit G, Erman M, Wang-Weigand S, Sainati S, Zhang J, Roth T. Evaluation of the efficacy and safety of ramelteon in subjects with chronic insomnia. *J Clin Sleep Med.* 2007;3(5):495–504.
122. Erman M, Seiden D, Zammit G, Sainati S, Zhang J. An efficacy, safety, and dose-response study of Ramelteon in patients with chronic primary insomnia. *Sleep Med.* 2006;7(1):17–24.
123. Mini L, Wang-Weigand S, Zhang J. Ramelteon 8 mg/d versus placebo in patients with chronic insomnia: post hoc analysis of a 5-week trial using 50% or greater reduction in latency to persistent sleep as a measure of treatment effect. *Clin Ther.* 2008;30(7):1316–1323.
124. Mini LJ, Wang-Weigand S, Zhang J. Self-reported efficacy and tolerability of ramelteon 8 mg in older adults experiencing severe sleep-onset difficulty. *Am J Geriatr Pharmacother.* 2007;5(3):177–184.
125. Roth T, Seiden D, Sainati S, Wang-Weigand S, Zhang J, Zee P. Effects of ramelteon on patient-reported sleep latency in older adults with chronic insomnia. *Sleep Med.* 2006;7(4):312–318.
126. Wang-Weigand S, Watissee M, Roth T. Use of a post-sleep questionnaire-interactive voice response system (PSQ-IVRS) to evaluate the subjective sleep effects of ramelteon in adults with chronic insomnia. *Sleep Med.* 2011;12(9):920–923.
127. Krystal AD, Durrence HH, Scharf M, et al. Efficacy and Safety of Doxepin 1 mg and 3 mg in a 12-week Sleep Laboratory and Outpatient Trial of Elderly Subjects with Chronic Primary Insomnia. *Sleep.* 2010;33(11):1553–1561.
128. Krystal AD, Lankford A, Durrence HH, et al. Efficacy and safety of doxepin 3 and 6 mg in a 35-day sleep laboratory trial in adults with chronic primary insomnia. *Sleep.* 2011;34(10):1433–1442.
129. Roth T, Rogowski R, Hull S, et al. Efficacy and safety of doxepin 1 mg, 3 mg, and 6 mg in adults with primary insomnia. *Sleep.* 2007;30(11):1555–1561.
130. Scharf M, Rogowski R, Hull S, et al. Efficacy and safety of doxepin 1 mg, 3 mg, and 6 mg in elderly patients with primary insomnia: a randomized, double-blind, placebo-controlled crossover study. *J Clin Psychiatry.* 2008;69(10):1557–1564.
131. Lankford A, Rogowski R, Essink B, Ludington E, Heith Durrence H, Roth T. Efficacy and safety of doxepin 6 mg in a four-week outpatient trial of elderly adults with chronic primary insomnia. *Sleep Med.* 2012;13(2):133–138.
132. Hajak G, Rodenbeck A, Voderholzer U, et al. Doxepin in the treatment of primary insomnia: a placebo-controlled, double-blind, polysomnographic study. *J Clin Psychiatry.* 2001;62(6):453–463.
133. Roth T, Wright KP, Jr., Walsh J. Effect of tiagabine on sleep in elderly subjects with primary insomnia: a randomized, double-blind, placebo-controlled study. *Sleep.* 2006;29(3):335–341.
134. Walsh JK, Perlis M, Rosenthal M, Krystal A, Jiang J, Roth T. Tiagabine increases slow-wave sleep in a dose-dependent fashion without affecting traditional efficacy measures in adults with primary insomnia. *J Clin Sleep Med.* 2006;2(1):35–41.
135. Walsh JK, Zammit G, Schweitzer PK, Ondrasik J, Roth T. Tiagabine enhances slow wave sleep and sleep maintenance in primary insomnia. *Sleep Med.* 2006;7(2):155–161.
136. Morin CM, Koetter U, Bastien C, Ware JC, Wooten V. Valerian-hops combination and diphenhydramine for treating insomnia: a randomized placebo-controlled clinical trial. *Sleep.* 2005;28(11):1465–1471.
137. Rickels K, Morris RJ, Newman H, Rosenfeld H, Schiller H, Weinstock R. Diphenhydramine in insomniac family practice patients: a double-blind study. *J Clin Pharmacol.* 1983;23(5-6):234–242.
138. Lemoine P, Nir T, Laudon M, Zisapel N. Prolonged-release melatonin improves sleep quality and morning alertness in insomnia patients aged 55 years and older and has no withdrawal effects. *J Sleep Res.* 2007;16(4):372–380.
139. Luthringer R, Muzet M, Zisapel N, Staner L. The effect of prolonged-release melatonin on sleep measures and psychomotor performance in elderly patients with insomnia. *Int Clin Psychopharmacol.* 2009;24(5):239–249.
140. Wade AG, Ford I, Crawford G, et al. Efficacy of prolonged release melatonin in insomnia patients aged 55-80 years: quality of sleep and next-day alertness outcomes. *Curr Med Res Opin.* 2007;23(10):2597–2605.
141. Baskett JJ, Broad JB, Wood PC, et al. Does melatonin improve sleep in older people? A randomised crossover trial. *Age Ageing.* 2003;32(2):164–170.
142. Garzon C, Guerrero JM, Aramburu O, Guzman T. Effect of melatonin administration on sleep, behavioral disorders and hypnotic drug discontinuation in the elderly: a randomized, double-blind, placebo-controlled study. *Aging Clin Exp Res.* 2009;21(1):38–42.
143. Haimov I, Lavie P, Laudon M, Herer P, Vigder C, Zisapel N. Melatonin replacement therapy of elderly insomniacs. *Sleep.* 1995;18(7):598–603.
144. Rondanelli M, Opizzi A, Monteferrario F, Antoniello N, Manni R, Klersy C. The effect of melatonin, magnesium, and zinc on primary insomnia in long-term care facility residents in Italy: a double-blind, placebo-controlled clinical trial. *J Am Geriatr Soc.* 2011;59(1):82–90.
145. Wade AG, Crawford G, Ford I, et al. Prolonged release melatonin in the treatment of primary insomnia: evaluation of the age cut-off for short- and long-term response. *Curr Med Res Opin.* 2011;27(1):87–98.
146. Wade AG, Ford I, Crawford G, et al. Nightly treatment of primary insomnia with prolonged release melatonin for 6 months: a randomized placebo controlled trial on age and endogenous melatonin as predictors of efficacy and safety. *BMC Med.* 2010;8:51.
147. Zhdanova IV, Wurtman RJ, Regan MM, Taylor JA, Shi JP, Leclair OU. Melatonin treatment for age-related insomnia. *J Clin Endocrinol Metab.* 2001;86(10):4727–4730.
148. Hudson C, Hudson SP, Hecht T, MacKenzie J. Protein source tryptophan versus pharmaceutical grade tryptophan as an efficacious treatment for chronic insomnia. *Nutr Neurosci.* 2005;8(2):121–127.
149. Hartmann E, Lindsley JG, Spinweber C. Chronic insomnia: effects of tryptophan, flurazepam, secobarbital, and placebo. *Psychopharmacology (Berl).* 1983;80(2):138–142.

150. Spinweber CL. L-tryptophan administered to chronic sleep-onset insomniacs: late-appearing reduction of sleep latency. *Psychopharmacology (Berl).* 1986;90(2):151–155.
151. Oxman AD, Flottorp S, Havelsrud K, et al. A televised, web-based randomised trial of an herbal remedy (valerian) for insomnia. *PLoS One.* 2007;2(10):e1040.
152. Cohn JB, Wilcox CS, Bremner J, Ettinger M. Hypnotic efficacy of estazolam compared with flurazepam in outpatients with insomnia. *J Clin Pharmacol.* 1991;31(8):747–750.
153. Dominguez RA, Goldstein BJ, Jacobson AF, Steinbook RM. Comparative efficacy of estazolam, flurazepam, and placebo in outpatients with insomnia. *J Clin Psychiatry.* 1986;47(7):362–365.
154. Scharf MB, Roth PB, Dominguez RA, Ware JC. Estazolam and flurazepam: a multicenter, placebo-controlled comparative study in outpatients with insomnia. *J Clin Pharmacol.* 1990;30(5):461–467.
155. Aden GC, Thatcher C. Quazepam in the short-term treatment of insomnia in outpatients. *J Clin Psychiatry.* 1983;44(12):454–456.
156. Hernandez Lara R, Del Rosal PL, Ponce MC. Short-term study of quazepam 15 milligrams in the treatment of insomnia. *J Int Med Res.* 1983;11(3):162–166.
157. Martinez HT, Serna CT. Short-term treatment with quazepam of insomnia in geriatric patients. *Clin Ther.* 1982;5(2):174–178.
158. Mendels J, Stern S. Evaluation of the short-term treatment of insomnia in outpatients with 15 milligrams of quazepam. *J Int Med Res.* 1983;11(3):155–161.
159. O'Hair DE, Winsauer HJ. Evaluation of short-term treatment with 30 mg of quazepam in insomniac outpatients. *Clin Ther.* 1981;4(4):291–301.
160. Roth TG, Roehrs TA, Koshorek GL, Greenblatt DJ, Rosenthal LD. Hypnotic effects of low doses of quazepam in older insomniacs. *J Clin Psychopharmacol.* 1997;17(5):401–406.
161. Elie R, Caille G, Levasseur FA, Gareau J. Comparative hypnotic activity of single doses of loprazolam, flurazepam, and placebo. *J Clin Pharmacol.* 1983;23(1):32–36.
162. Elie R, Lavoie G, Bourgouin J, Le Morvan P. Zopiclone versus flurazepam in insomnia: prolonged administration and withdrawal. *Int Clin Psychopharmacol.* 1990;5(4):279–286.
163. Kripke DF, Hauri P, Ancoli-Israel S, Roth T. Sleep evaluation in chronic insomniacs during 14-day use of flurazepam and midazolam. *J Clin Psychopharmacol.* 1990;10(4 Suppl):32S–43S.
164. Mamelak M, Buck L, Csima A, Price V, Smiley A. Effects of flurazepam and zopiclone on the performance of chronic insomniac patients: a study of ethanol-drug interaction. *Sleep.* 1987;10 Suppl 1:79–87.
165. Mamelak M, Csima A, Buck L, Price V. A comparative study on the effects of brotizolam and flurazepam on sleep and performance in the elderly. *J Clin Psychopharmacol.* 1989;9(4):260–267.
166. Melo de Paula AJ. Comparative study of lormetazepam and flurazepam in the treatment of insomnia. *Clin Ther.* 1984;6(4):500–508.
167. Salkind MR, Silverstone T. A clinical and psychometric evaluation of flurazepam. *Br J Clin Pharmacol.* 1975;2(3):223–226.
168. Gotestam KG, Oppoyen F, Berntzen D. Treatment of insomnia with two benzodiazepines: a double-blind crossover study. *Eur J Clin Pharmacol.* 1991;41(2):137–140.
169. Tassniyom K, Paholpak S, Tassniyom S, Kiewyoo J. Quetiapine for primary insomnia: a double blind, randomized controlled trial. *J Med Assoc Thai.* 2010;93(6):729–734.
170. Lo HS, Yang CM, Lo HG, Lee CY, Ting H, Tzang BS. Treatment effects of gabapentin for primary insomnia. *Clin Neuropharmacol.* 2010;33(2):84–90.
171. Nowell PD, Reynolds CF, 3rd, Buysse DJ, Dew MA, Kupfer DJ. Paroxetine in the treatment of primary insomnia: preliminary clinical and electroencephalogram sleep data. *J Clin Psychiatry.* 1999;60(2):89–95.
172. Reynolds CF, 3rd, Buysse DJ, Miller MD, Pollock BG, Hall M, Mazumdar S. Paroxetine treatment of primary insomnia in older adults. *Am J Geriatr Psychiatry.* 2006;14(9):803–807.
173. Hohagen F, Montero RF, Weiss E, et al. Treatment of primary insomnia with trimipramine: an alternative to benzodiazepine hypnotics? *Eur Arch Psychiatry Clin Neurosci.* 1994;244(2):65–72.
174. Riemann D, Voderholzer U, Cohrs S, et al. Trimipramine in primary insomnia: results of a polysomnographic double-blind controlled study. *Pharmacopsychiatry.* 2002;35(5):165–174.
175. Smith MT, Perlis ML, Park A, et al. Comparative meta-analysis of pharmacotherapy and behavior therapy for persistent insomnia. *Am J Psychiatry.* 2002;159(1):5–11.

ACKNOWLEDGMENTS

The task force thanks and acknowledges the contributions of Karl Doghramji, MD, who served as a critical reviewer of this guideline.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication September, 2016
Submitted in final revised form September, 2016
Accepted for publication September, 2016
Address correspondence to: Michael J. Sateia, MD, Geisel School of Medicine at Dartmouth, Hanover, NH 03755; Tel: (603) 650-7534; Fax: (603) 650-7820; Email: research@aasmnet.org

DISCLOSURE STATEMENT

The development of this clinical practice guideline was funded by the American Academy of Sleep Medicine. Dr. Neubauer is a member of the Board of Directors for the National Sleep Foundation; and he has been a consultant for Purdue Pharama. Dr. Krystal serves on a scientific advisory board for Merck, and therefore did not participate in the development of the suvorexant recommendation; he has received research support from the NIH, TEVA and Sunovion; and he has been a consultant for Flamel, Atentiv, Ostuka, Neurocrine, Lundbeck, Pernix, Janssen, Jazz and Merck. Dr. Buysse has been a consultant for Cereve, Inc, Emmi Solutions, Philips Respironics, BeHealth; he has received research support from the NIH; and he owns intellectual property rights in the Pittsburgh Sleep Quality Index (PSQI). Mr. Heald is employed by the American Academy of Sleep Medicine. The other authors have indicated no financial conflicts of interest.

Literature Search Terms:

((insomnia OR "Sleep Initiation and Maintenance Disorders"[Mesh]) AND ((clinical trial) OR (randomized controlled trial)) NOT "Editorial"[Publication Type] NOT "Letter"[Publication Type] NOT "Comment"[Publication Type] NOT "Case Reports"[Publication Type] NOT "Biography"[Publication Type] NOT "Review"[Publication Type] NOT (transient[TI])) NOT (animals[mh] NOT humans[mh])

Suvorexant - Summary of Findings Tables

Table S1 – Summary of Findings table for suvorexant 10 mg for the treatment of chronic insomnia

References: Herring 2012(A)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 10 mg Suvorexant vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean sleep latency in the suvorexant group was **2.3 minutes lower** (13.68 lower to 9.08 higher)	175 (1 study)[A]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[2,3]	The mean wake after sleep onset in the suvorexant group was **21.5 minutes lower** (36.34 to 6.66 lower)	175 (1 study)[A]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[2,4]	The mean sleep efficiency in the suvorexant group was **4.7 percent higher** (0.97 to 8.43 higher)	175 (1 study)[A]

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-13.68 to 9.08) crosses the Clinical Significance Threshold (10 min)
[2] Study funded by industry
[3] 95% CI (-36.34 to -6.66) crosses the Clinical Significance Threshold (20 min)
[4] 95% CI (0.97 to 8.43) crosses the Clinical Significance Threshold (5%)

Table S2 – Summary of Findings table for suvorexant 15/20 mg for the treatment of chronic insomnia

References: Herring 2016(A)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 15/20 mg Suvorexant vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean sleep latency in the suvorexant group was **8.1 minutes lower** (13.85 to 2.35 lower)	423 (1 study)[A]
Sleep Latency (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean sleep latency in the suvorexant group was **5.2 minutes lower** (10.1 to 0.3 lower)	567 (1 study)[A]
Total Sleep Time* (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean total sleep time in the suvorexant group was **10.6 minutes higher** (1.79 to 19.41 higher)	567 (1 study)[A]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[2,3]	The mean wake after sleep onset in the suvorexant group was **16.60 minutes lower** (24.87 to 8.33 lower)	567 (1 study)[A]

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-13.85 to -2.35) crosses Clinical Significance Threshold (10 min)
[2] Study funded by industry
[3] 95% CI (-24.87 to -8.33) crosses Clinical Significance Threshold (20 min)

Table S3 – Summary of Findings table for suvorexant 20 mg for the treatment of chronic insomnia

References: Herring 2012(A)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 20 mg Suvorexant vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊕⊖ **moderate**[2]	The mean sleep latency in the suvorexant group was **22.3 minutes lower** (33.77 to 10.83 lower)	173 (1 study)[A]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean wake after sleep onset in the suvorexant group was **28.1 minutes lower** (43.07 to 13.13 lower)	173 (1 study)[A]
Sleep Efficiency (PSG)	⊕⊕⊕⊖ **moderate**[2]	The mean sleep efficiency in the suvorexant group was **10.4 percent higher** (6.65 to 14.15 higher)	173 (1 study)[A]

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-43.07 to -13.13) crosses Clinical Significance Threshold (20 min)
[2] Study funded by industry

Eszopiclone - Meta-Analyses and Summary of Findings Tables

Figure S1 – Meta-analysis of data for PSG-determined sleep latency in response to eszopiclone 2 mg

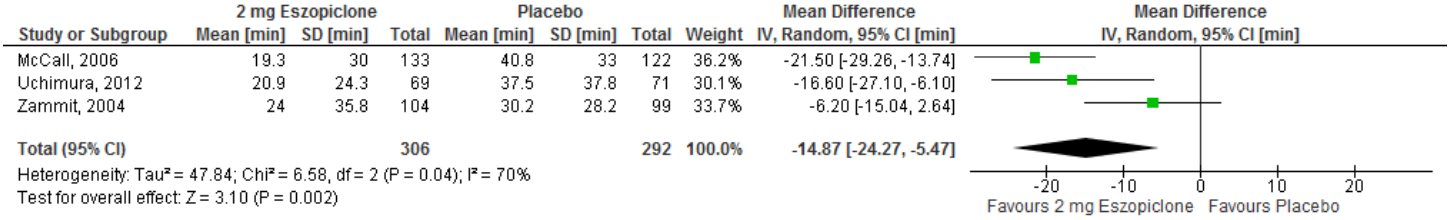

2 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
McCall, 2006 19.3 30 133 40.8 33 122 36.2% -21.50 [-29.26, -13.74]
Uchimura, 2012 20.9 24.3 69 37.5 37.8 71 30.1% -16.60 [-27.10, -6.10]
Zammit, 2004 24 35.8 104 30.2 28.2 99 33.7% -6.20 [-15.04, 2.64]
Total (95% CI) 306 292 100.0% -14.87 [-24.27, -5.47]
Heterogeneity: Tau² = 47.84; Chi² = 6.58, df = 2 (P = 0.04); I² = 70%
Test for overall effect: Z = 3.10 (P = 0.002)
-20 -10 0 10 20
Favours 2 mg Eszopiclone
Favours Placebo

Figure S2 – Meta-analysis of data for subjectively-determined sleep latency in response to eszopiclone 2 mg

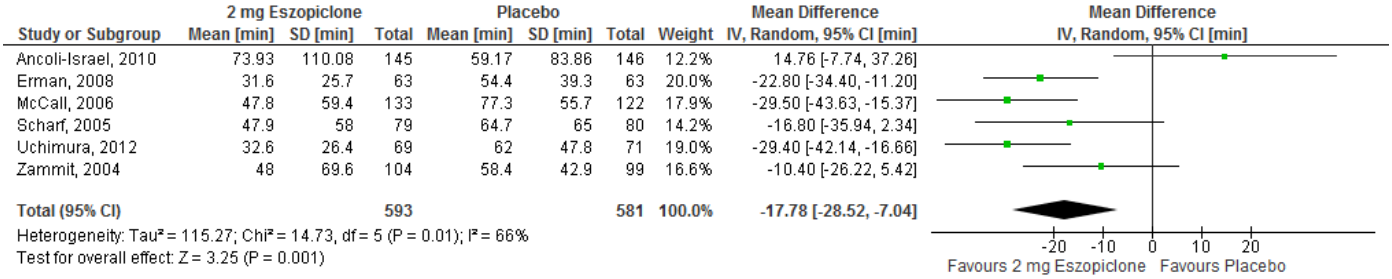


Figure S3 – Meta-analysis of data for subjectively-determined total sleep time in response to eszopiclone 2 mg

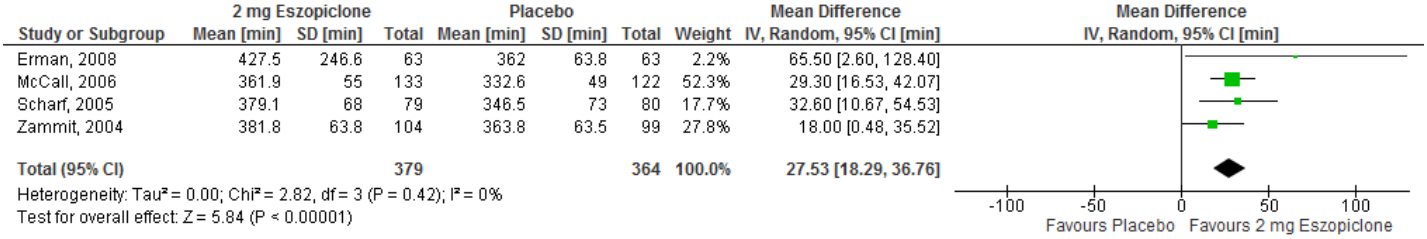

2 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Erman, 2008 427.5 246.6 63 362 63.8 63 2.2% 65.50 [2.60, 128.40]
McCall, 2006 361.9 55 133 332.6 49 122 52.3% 29.30 [16.53, 42.07]
Scharf, 2005 379.1 68 79 346.5 73 80 17.7% 32.60 [10.67, 54.53]
Zammit, 2004 381.8 63.8 104 363.8 63.5 99 27.8% 18.00 [0.48, 35.52]
Total (95% CI) 379 364 100.0% 27.53 [18.29, 36.76]
Heterogeneity: Tau² = 0.00; Chi² = 2.82, df = 3 (P = 0.42); I² = 0%
Test for overall effect: Z = 5.84 (P < 0.00001)
-100 -50 0 50 100
Favours Placebo
Favours 2 mg Eszopiclone

Figure S4 – Meta-analysis of data for PSG-determined wake after sleep onset in response to eszopiclone 2 mg

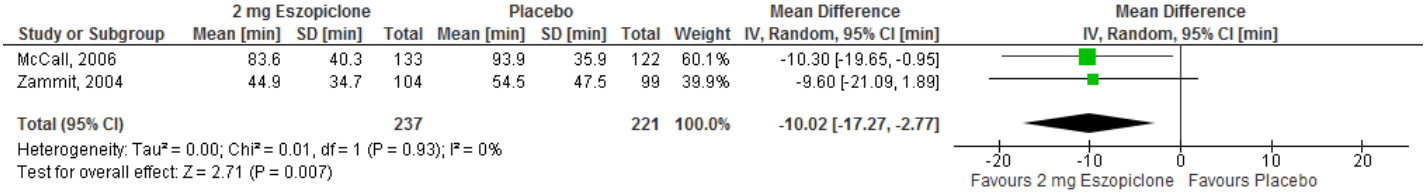

2 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
McCall, 2006 83.6 40.3 133 93.9 35.9 122 60.1% -10.30 [-19.65, -0.95]
Zammit, 2004 44.9 34.7 104 54.5 47.5 99 39.9% -9.60 [-21.09, 1.89]
Total (95% CI) 237 221 100.0% -10.02 [-17.27, -2.77]
Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.93); I² = 0%
Test for overall effect: Z = 2.71 (P = 0.007)
-20 -10 0 10 20
Favours 2 mg Eszopiclone
Favours Placebo

Figure S5 – Meta-analysis of data for subjectively-determined wake after sleep onset in response to eszopiclone 2 mg

Study or Subgroup	2 mg Eszopiclone Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Ancoli-Israel, 2010	63.58	63.05	145	68.05	54.63	146	23.6%	-4.47 [-18.03, 9.09]
Erman, 2008	69	229.8	63	56.6	48.3	63	1.5%	12.40 [-45.58, 70.38]
McCall, 2006	77.3	45.7	133	88.1	48.3	122	30.7%	-10.80 [-22.37, 0.77]
Scharf, 2005	54.1	61	79	67.4	56	80	14.0%	-13.30 [-31.51, 4.91]
Zammit, 2004	53.4	48.1	104	49.1	36.1	99	30.3%	4.30 [-7.36, 15.96]
Total (95% CI)			**524**			**510**	**100.0%**	**-4.74 [-11.87, 2.39]**

Heterogeneity: Tau² = 8.26; Chi² = 4.55, df = 4 (P = 0.34); I² = 12%
Test for overall effect: Z = 1.30 (P = 0.19)

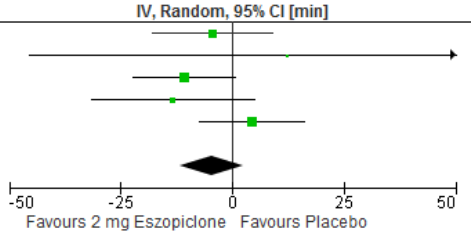


Figure S6 – Meta-analysis of data for subjectively-determined quality of sleep in response to eszopiclone 2 mg

Study or Subgroup	2 mg Eszopiclone Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Erman, 2008	57.2	20.9	63	44.4	19.9	63	19.7%	0.62 [0.27, 0.98]
Scharf, 2005	7.5	1.7	79	6.6	1.9	80	25.3%	0.50 [0.18, 0.81]
Uchimura, 2012	6.3	1.8	69	5.2	1.9	71	22.0%	0.59 [0.25, 0.93]
Zammit, 2004	54.4	18.7	104	49	18.1	99	33.0%	0.29 [0.02, 0.57]
Total (95% CI)			**315**			**313**	**100.0%**	**0.47 [0.32, 0.63]**

Heterogeneity: Tau² = 0.00; Chi² = 2.81, df = 3 (P = 0.42); I² = 0%
Test for overall effect: Z = 5.86 (P < 0.00001)

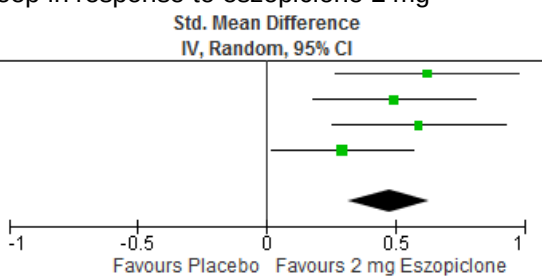


Figure S7 – Meta-analysis of data for PSG-determined sleep efficiency in response to eszopiclone 2 mg

Study or Subgroup	2 mg Eszopiclone Mean [%]	SD [%]	Total	Placebo Mean [%]	SD [%]	Total	Weight	Mean Difference IV, Random, 95% CI [%]
McCall, 2006	79.4	9	133	73.4	9	122	56.8%	6.00 [3.79, 8.21]
Zammit, 2004	86.2	9.6	104	82.9	11.7	99	43.2%	3.30 [0.35, 6.25]
Total (95% CI)			**237**			**221**	**100.0%**	**4.83 [2.21, 7.46]**

Heterogeneity: Tau² = 1.87; Chi² = 2.06, df = 1 (P = 0.15); I² = 51%
Test for overall effect: Z = 3.61 (P = 0.0003)

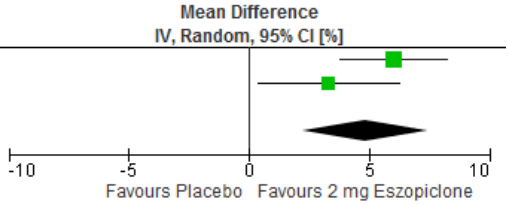


Figure S8 – Meta-analysis of data for PSG-determined number of awakenings in response to eszopiclone 2 mg

Study or Subgroup	2 mg Eszopiclone Mean [#]	SD [#]	Total	Placebo Mean [#]	SD [#]	Total	Weight	Mean Difference IV, Random, 95% CI [#]
McCall, 2006	9.1	3.3	133	9.5	3.3	122	56.5%	-0.40 [-1.21, 0.41]
Zammit, 2004	7.3	4	104	6.5	4.5	99	43.5%	0.80 [-0.37, 1.97]
Total (95% CI)			**237**			**221**	**100.0%**	**0.12 [-1.04, 1.29]**

Heterogeneity: Tau² = 0.46; Chi² = 2.72, df = 1 (P = 0.10); I² = 63%
Test for overall effect: Z = 0.21 (P = 0.84)

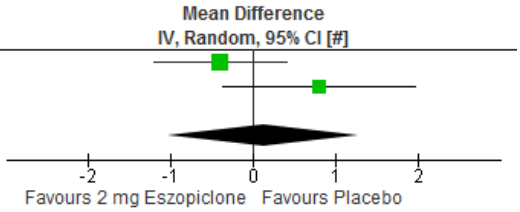


Figure S9 – Meta-analysis of data for subjectively-determined number of awakenings in response to eszopiclone 2 mg

Study or Subgroup	2 mg Eszopiclone Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Erman, 2008	3	1.4	63	3.7	2.1	63	7.9%	-0.70 [-1.32, -0.08]
McCall, 2006	2	1	133	2.3	1	122	50.9%	-0.30 [-0.55, -0.05]
Scharf, 2005	1.5	1.1	79	1.8	1	80	28.8%	-0.30 [-0.63, 0.03]
Zammit, 2004	2.9	1.7	104	3.2	1.9	99	12.4%	-0.30 [-0.80, 0.20]
Total (95% CI)			**379**			**364**	**100.0%**	**-0.33 [-0.51, -0.16]**

Heterogeneity: Tau² = 0.00; Chi² = 1.46, df = 3 (P = 0.69); I² = 0%
Test for overall effect: Z = 3.71 (P = 0.0002)

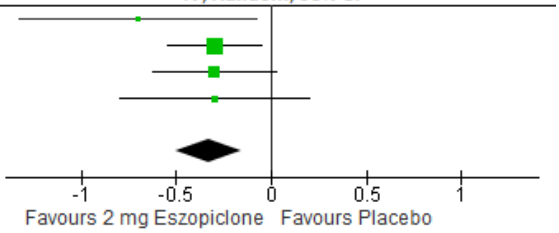


Figure S10– Meta-analysis of data for the occurrence of dizziness in response to eszopiclone 2 mg

Study or Subgroup	Eszopiclone 2 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Erman 2008	0	63	3	63	21.5%	-0.05 [-0.11, 0.01]
McCall 2006	12	133	2	122	23.2%	0.07 [0.02, 0.13]
Uchimura 2012	3	104	4	99	24.2%	-0.01 [-0.06, 0.04]
Zammit 1971	0	69	0	71	31.1%	0.00 [-0.03, 0.03]
Total (95% CI)		**369**		**355**	**100.0%**	**0.00 [-0.04, 0.05]**
Total events	15		9			

Heterogeneity: Tau² = 0.00; Chi² = 10.93, df = 3 (P = 0.01); I² = 73%
Test for overall effect: Z = 0.18 (P = 0.85)

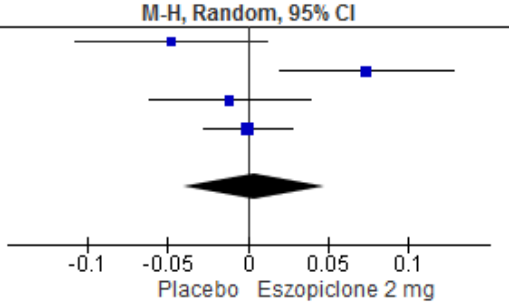


Figure S11– Meta-analysis of data for the occurrence of dry mouth in response to eszopiclone 2 mg

Study or Subgroup	Eszopiclone 2 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
McCall 2006	16	133	2	122	47.7%	0.10 [0.04, 0.16]
Zammit 1971	5	104	2	99	52.3%	0.03 [-0.02, 0.08]
Total (95% CI)		**237**		**221**	**100.0%**	**0.06 [-0.01, 0.14]**
Total events	21		4			

Heterogeneity: Tau² = 0.00; Chi² = 4.02, df = 1 (P = 0.04); I² = 75%
Test for overall effect: Z = 1.62 (P = 0.11)

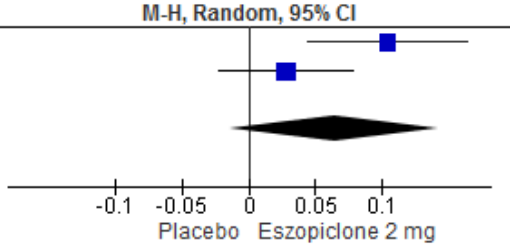


Figure S12– Meta-analysis of data for the occurrence of headache in response to eszopiclone 2 mg

Study or Subgroup	Eszopiclone 2 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Ancoli-Israel 2010	20	145	18	146	33.4%	0.01 [-0.06, 0.09]
Erman 2008	4	63	6	63	22.6%	-0.03 [-0.13, 0.06]
Scharf 2005	12	79	12	80	16.2%	0.00 [-0.11, 0.11]
Zammit 1971	14	104	8	99	27.9%	0.05 [-0.03, 0.14]
Total (95% CI)		**391**		**388**	**100.0%**	**0.01 [-0.03, 0.06]**
Total events	50		44			

Heterogeneity: Tau² = 0.00; Chi² = 1.80, df = 3 (P = 0.61); I² = 0%
Test for overall effect: Z = 0.57 (P = 0.57)

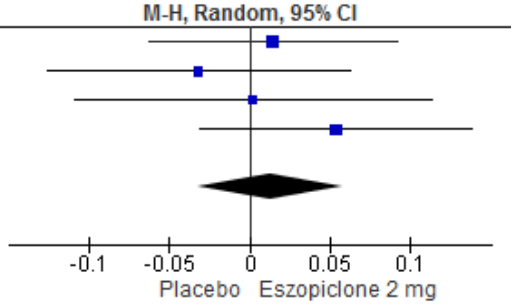


Figure S13– Meta-analysis of data for the occurrence of somnolence in response to eszopiclone 2 mg

Study or Subgroup	Eszopiclone 2 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Erman 2008	2	63	2	63	19.6%	0.00 [-0.06, 0.06]
McCall 2006	12	133	9	122	16.2%	0.02 [-0.05, 0.08]
Scharf 2005	3	79	7	80	13.1%	-0.05 [-0.12, 0.03]
Uchimura 2012	8	104	3	99	19.5%	0.05 [-0.01, 0.11]
Zammit 1971	2	69	1	71	31.6%	0.01 [-0.03, 0.06]
Total (95% CI)		**448**		**435**	**100.0%**	**0.01 [-0.02, 0.04]**
Total events	27		22			

Heterogeneity: Tau² = 0.00; Chi² = 3.98, df = 4 (P = 0.41); I² = 0%
Test for overall effect: Z = 0.72 (P = 0.47)

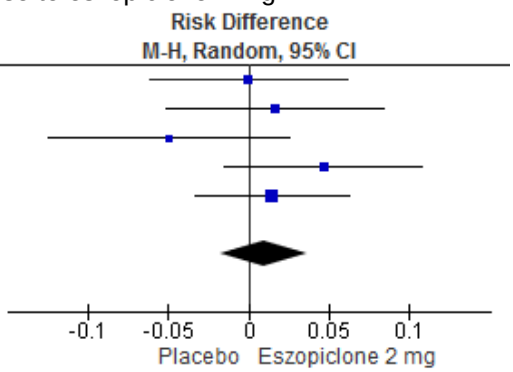


Figure S14– Meta-analysis of data for the occurrence of unpleasant taste in response to eszopiclone 2 mg

Study or Subgroup	Eszopiclone 2 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Ancoli-Israel 2010	18	145	2	146	21.2%	0.11 [0.05, 0.17]
Erman 2008	3	63	1	63	20.6%	0.03 [-0.03, 0.09]
McCall 2006	16	133	0	122	21.2%	0.12 [0.06, 0.18]
Scharf 2005	3	79	7	80	18.9%	-0.05 [-0.12, 0.03]
Zammit 1971	18	104	3	99	18.2%	0.14 [0.06, 0.22]
Total (95% CI)		**524**		**510**	**100.0%**	**0.07 [0.01, 0.14]**
Total events	58		13			

Heterogeneity: Tau² = 0.00; Chi² = 19.57, df = 4 (P = 0.0006); I² = 80%
Test for overall effect: Z = 2.20 (P = 0.03)

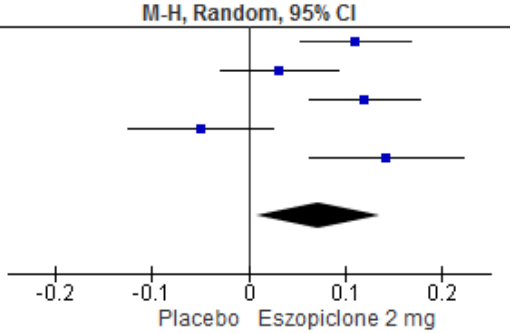


Table S4 – Summary of Findings table for eszopiclone 2 mg for the treatment of chronic insomnia

References: Ancoli-Israel 2010(A); Erman 2008(B); McCall 2006(C); Scharf 2005(D); Uchimura 2012(E); Zammit 2004(F)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 2 mg Eszopiclone vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean sleep latency in the eszopiclone groups was **14.87 minutes lower** (24.27 to 5.47 lower)	598 (3 studies)[C,E,F]
Sleep Latency (Subjective)	⊕⊕⊖⊖ **low**[2,3]	The mean sleep latency in the eszopiclone groups was **17.78 minutes lower** (28.52 to 7.04 lower)	1174 (6 studies)[A,B,C,D,E,F]
Total Sleep Time* (Subjective)	⊕⊕⊖⊖ **low**[2,4]	The mean total sleep time in the eszopiclone groups was **27.53 minutes higher** (18.29 to 36.76 higher)	743 (4 studies)[B,C,D,F]
Wake After Sleep Onset* (PSG)	⊕⊕⊕⊖ **moderate**[2]	The mean wake after sleep onset in the eszopiclone groups was **10.02 minutes lower** (17.27 to 2.77 lower)	458 (2 studies)[C,F]
Wake After Sleep Onset (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean wake after sleep onset in the eszopiclone groups was **4.74 minutes lower** (11.87 lower to 2.39 higher)	1034 (5 studies)[A,B,C,D,F]
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[2,6]	The mean quality of sleep in the eszopiclone groups was **0.47 standard deviations higher** (0.32 to 0.63 higher)	628 (4 studies)[B,D,E,F]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[2,5]	The mean sleep efficiency in the eszopiclone groups was **4.83 percent higher** (2.21 to 7.46 higher)	458 (2 studies)[C,F]
Sleep Efficiency (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean sleep efficiency in the eszopiclone groups was **0.30 percent lower** (0.79 lower to 0.19 higher)	203 (1 study)[F]
Number of Awakening (PSG)	⊕⊕⊕⊖ **moderate**[2]	The mean number awakening in the eszopiclone groups was **0.12 awakenings higher** (1.04 lower to 1.29 higher)	458 (2 studies)[C,F]
Number of Awakenings (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean number of awakenings in the eszopiclone groups was **0.33 awakenings lower** (0.51 to 0.16 lower)	743 (4 studies)[B,C,D,F]

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-24.27, -5.47) crosses Clinical Signficance (10 min)
[2] All studies funded by industry
[3] 95% CI (-33.81, -6.35) crosses Clinical Signficance (20 min)
[4] 95% CI (18.29, 36.76) crosses Clinical Signficance (20 min)
[5] 95% CI (2.21, 7.46) crosses Clinical Significance (5%)
[6] 95% CI (0.37, 0.76) crosses Clinical Significance (SMD 0.5)

Figure S15 – Meta-analysis of data for PSG-determined sleep latency in response to eszopiclone 3 mg

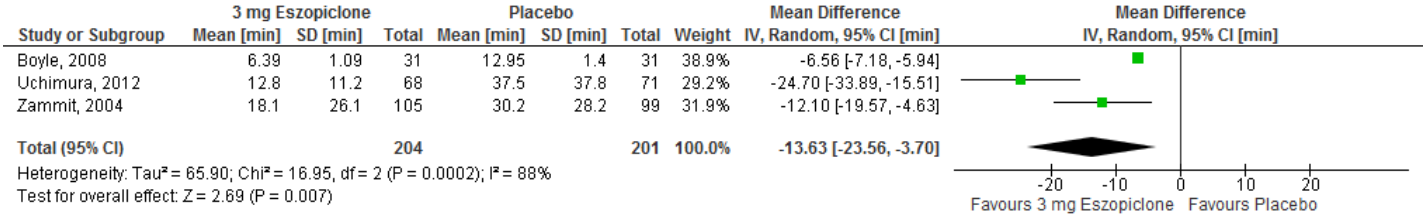

3 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Boyle, 2008 6.39 1.09 31 12.95 1.4 31 38.9% -6.56 [-7.18, -5.94]
Uchimura, 2012 12.8 11.2 68 37.5 37.8 71 29.2% -24.70 [-33.89, -15.51]
Zammit, 2004 18.1 26.1 105 30.2 28.2 99 31.9% -12.10 [-19.57, -4.63]
Total (95% CI) 204 201 100.0% -13.63 [-23.56, -3.70]
Heterogeneity: Tau² = 65.90; Chi² = 16.95, df = 2 (P = 0.0002); I² = 88%
Test for overall effect: Z = 2.69 (P = 0.007)
-20 -10 0 10 20
Favours 3 mg Eszopiclone Favours Placebo

Figure S16 – Meta-analysis of data for subjectively-determined sleep latency in response to eszopiclone 3 mg

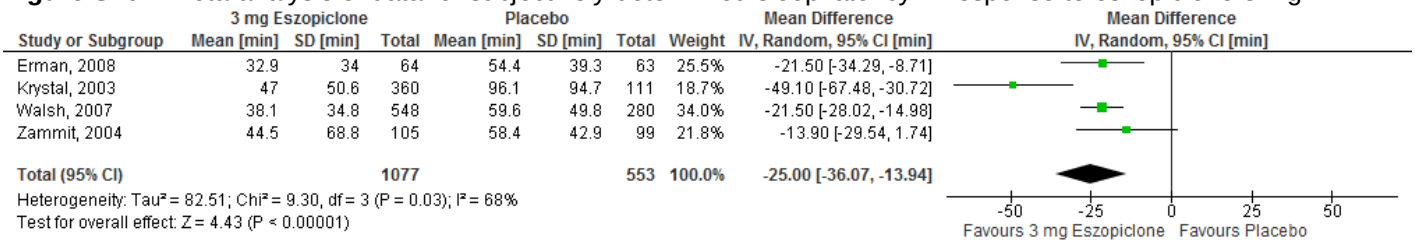

3 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Erman, 2008 32.9 34 64 54.4 39.3 63 25.5% -21.50 [-34.29, -8.71]
Krystal, 2003 47 50.6 360 96.1 94.7 111 18.7% -49.10 [-67.48, -30.72]
Walsh, 2007 38.1 34.8 548 59.6 49.8 280 34.0% -21.50 [-28.02, -14.98]
Zammit, 2004 44.5 68.8 105 58.4 42.9 99 21.8% -13.90 [-29.54, 1.74]
Total (95% CI) 1077 553 100.0% -25.00 [-36.07, -13.94]
Heterogeneity: Tau² = 82.51; Chi² = 9.30, df = 3 (P = 0.03); I² = 68%
Test for overall effect: Z = 4.43 (P < 0.00001)
-50 -25 0 25 50
Favours 3 mg Eszopiclone Favours Placebo

Figure S17 – Meta-analysis of data for subjectively-determined total sleep time in response to eszopiclone 3 mg

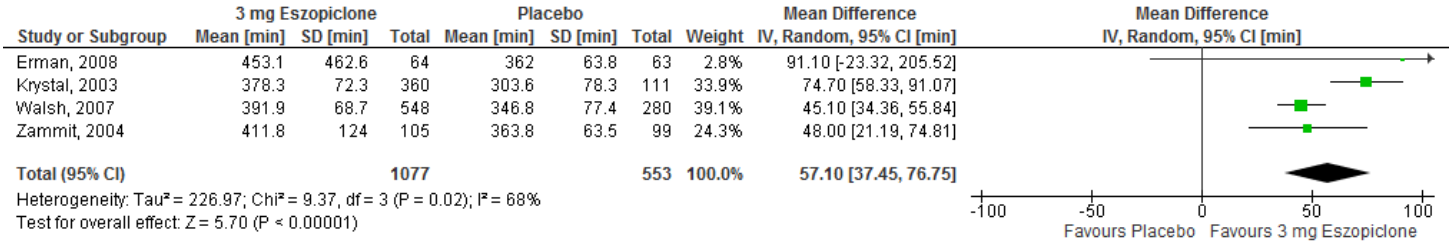

3 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Erman, 2008 453.1 462.6 64 362 63.8 63 2.8% 91.10 [-23.32, 205.52]
Krystal, 2003 378.3 72.3 360 303.6 78.3 111 33.9% 74.70 [58.33, 91.07]
Walsh, 2007 391.9 68.7 548 346.8 77.4 280 39.1% 45.10 [34.36, 55.84]
Zammit, 2004 411.8 124 105 363.8 63.5 99 24.3% 48.00 [21.19, 74.81]
Total (95% CI) 1077 553 100.0% 57.10 [37.45, 76.75]
Heterogeneity: Tau² = 226.97; Chi² = 9.37, df = 3 (P = 0.02); I² = 68%
Test for overall effect: Z = 5.70 (P < 0.00001)
-100 -50 0 50 100
Favours Placebo Favours 3 mg Eszopiclone

Figure S18 – Meta-analysis of data for PSG-determined wake after sleep onset in response to eszopiclone 3 mg

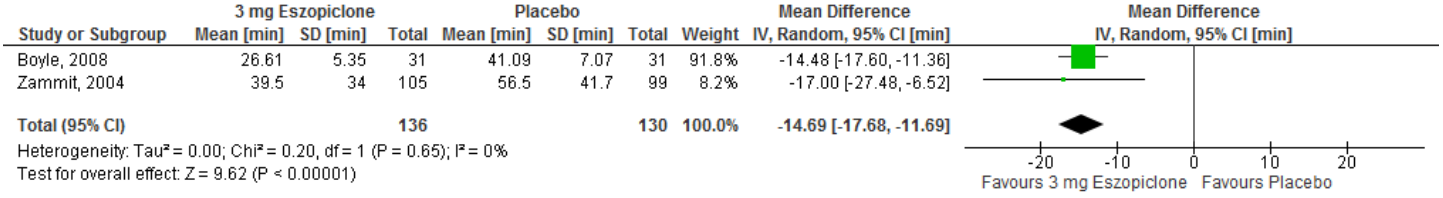

3 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Boyle, 2008 26.61 5.35 31 41.09 7.07 31 91.8% -14.48 [-17.60, -11.36]
Zammit, 2004 39.5 34 105 56.5 41.7 99 8.2% -17.00 [-27.48, -6.52]
Total (95% CI) 136 130 100.0% -14.69 [-17.68, -11.69]
Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 1 (P = 0.65); I² = 0%
Test for overall effect: Z = 9.62 (P < 0.00001)
-20 -10 0 10 20
Favours 3 mg Eszopiclone Favours Placebo

Figure S19 – Meta-analysis of data for subjectively-determined wake after sleep onset in response to eszopiclone 3 mg

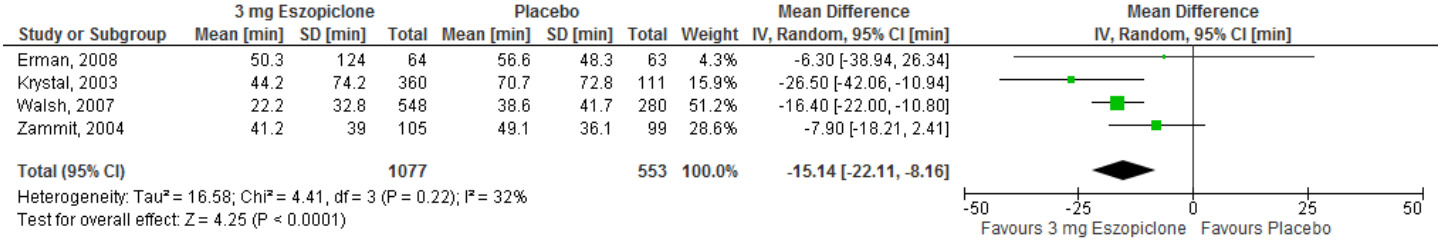

3 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Erman, 2008 50.3 124 64 56.6 48.3 63 4.3% -6.30 [-38.94, 26.34]
Krystal, 2003 44.2 74.2 360 70.7 72.8 111 15.9% -26.50 [-42.06, -10.94]
Walsh, 2007 22.2 32.8 548 38.6 41.7 280 51.2% -16.40 [-22.00, -10.80]
Zammit, 2004 41.2 39 105 49.1 36.1 99 28.6% -7.90 [-18.21, 2.41]
Total (95% CI) 1077 553 100.0% -15.14 [-22.11, -8.16]
Heterogeneity: Tau² = 16.58; Chi² = 4.41, df = 3 (P = 0.22); I² = 32%
Test for overall effect: Z = 4.25 (P < 0.0001)
-50 -25 0 25 50
Favours 3 mg Eszopiclone Favours Placebo

Figure S20 – Meta-analysis of data for subjectively-determined quality of sleep in response to eszopiclone 3 mg

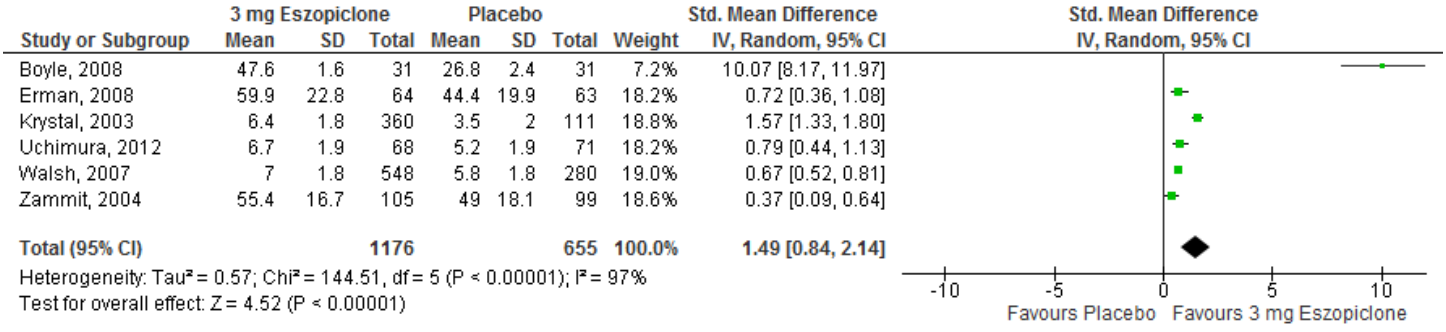


Figure S21 – Meta-analysis of data for PSG-determined sleep efficiency in response to eszopiclone 3 mg

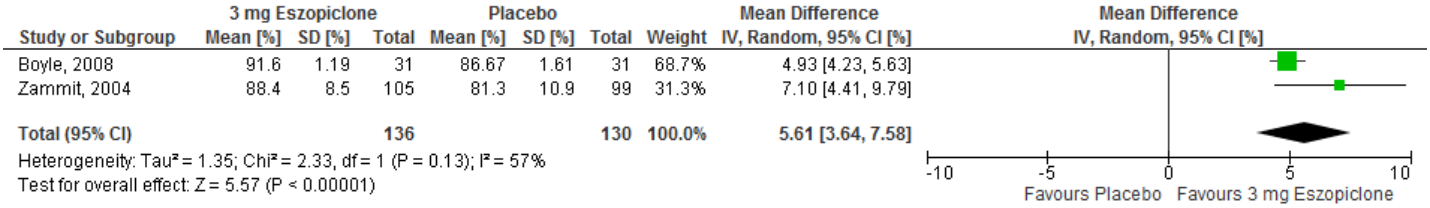

3 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [%] SD [%] Total Mean [%] SD [%] Total Weight IV, Random, 95% CI [%]
IV, Random, 95% CI [%]
Boyle, 2008 91.6 1.19 31 86.67 1.61 31 68.7% 4.93 [4.23, 5.63]
Zammit, 2004 88.4 8.5 105 81.3 10.9 99 31.3% 7.10 [4.41, 9.79]
Total (95% CI) 136 130 100.0% 5.61 [3.64, 7.58]
Heterogeneity: Tau² = 1.35; Chi² = 2.33, df = 1 (P = 0.13); I² = 57%
Test for overall effect: Z = 5.57 (P < 0.00001)
-10
-5
0
5
10
Favours Placebo
Favours 3 mg Eszopiclone

Figure S22 – Meta-analysis of data for subjectively-determined number of awakenings in response to eszopiclone 3 mg

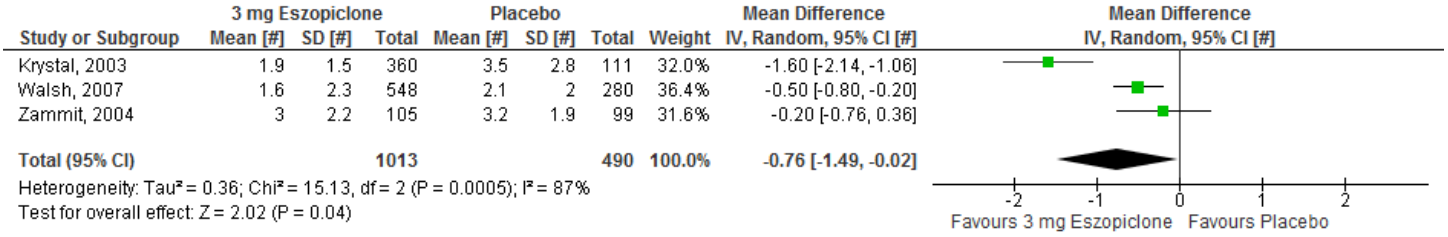

3 mg Eszopiclone
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [#] SD [#] Total Mean [#] SD [#] Total Weight IV, Random, 95% CI [#]
IV, Random, 95% CI [#]
Krystal, 2003 1.9 1.5 360 3.5 2.8 111 32.0% -1.60 [-2.14, -1.06]
Walsh, 2007 1.6 2.3 548 2.1 2 280 36.4% -0.50 [-0.80, -0.20]
Zammit, 2004 3 2.2 105 3.2 1.9 99 31.6% -0.20 [-0.76, 0.36]
Total (95% CI) 1013 490 100.0% -0.76 [-1.49, -0.02]
Heterogeneity: Tau² = 0.36; Chi² = 15.13, df = 2 (P = 0.0005); I² = 87%
Test for overall effect: Z = 2.02 (P = 0.04)
-2
-1
0
1
2
Favours 3 mg Eszopiclone
Favours Placebo

Table S5 – Summary of Findings table for eszopiclone 3 mg for the treatment of chronic insomnia

References: Boyle 2008(A); Erman 2008(B); Krystal 2003(C); Uchimura 2012(D); Walsh 2007(E); Zammit 2004(F)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 3 mg Eszopiclone vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊖⊖⊖ **very low**[1,2,3]	The mean sleep latency in the eszopiclone groups was **13.63 minutes lower** (23.56 to 3.7 lower)	405 (3 studies)[A,D,F]
Sleep Latency (Subjective)	⊕⊕⊖⊖ **low**[3,4]	The mean sleep latency in the eszopiclone groups was **25.00 minutes lower** (36.07 to 13.94 lower)	1630 (4 studies)[B,C,E,F]
Total Sleep Time* (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean total sleep time in the eszopiclone groups was **57.10 minutes higher** (37.45 to 76.75 higher)	1630 (4 studies)[B,C,E,F]
Wake After Sleep Onset* (PSG)	⊕⊕⊕⊖ **moderate**[3]	The mean wake after sleep onset in the eszopiclone groups was **14.69 minutes lower** (17.68 to 11.69 lower)	266 (2 studies)[A,F]
Wake After Sleep Onset (Subjective)	⊕⊕⊖⊖ **low**[3,5]	The mean wake after sleep onset in the eszopiclone groups was **15.14 minutes lower** (22.11 to 8.16 lower)	1630 (4 studies)[B,C,E,F]
Quality of Sleep* (Subjective)	⊕⊕⊖⊖ **low**[3,9]	The mean quality of sleep in the eszopiclone groups was **1.49 standard deviations higher** (0.84 to 2.14 higher)	1769 (6 studies)[A,B,C,D,E,F]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[3,6]	The mean sleep efficiency in the eszopiclone groups was **5.61 percent higher** (3.64 to 7.58 higher)	266 (2 studies)[A,F]
Number of Awakenings (Subjective)	⊕⊖⊖⊖ **very low**[3,6,7]	The mean number awakenings in the eszopiclone groups was **0.76 awakenings lower** (1.49 to 0.02 lower)	1503 (3 studies)[C,E,F]

* Critical Outcome, used to determine Quality of Evidence
[1] Heterogeneity (I^2 = 88%) greater than allowance (75%)
[2] 95% CI (-23.56, -3.70) crosses Clinical Significance (10 min)
[3] All studies funded by industry
[4] 95% CI (-36.07, -13.94) crosses Clinical Significance (20 min)
[5] 95% CI (-22.11, -8.16) crossess Clinical Significance (20 min)
[6] Heterogeneity (I^2 = 87%) greater than allowance (75%)
[7] 95% CI (-1.49, -0.02) crosses Clinical Significance (0.5 awakenings)
[8] 95% CI (3.64, 7.58) crosses Clinical Significance
[9] Heterogeneity (I^2 = 93%) greater than allowance (75%)

Zaleplon - Summary of Findings Tables

Table S6 – Summary of Findings table for zaleplon 5 mg for the treatment of chronic insomnia

References: Hedner 2000(A)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 5 mg Zaleplon vs Placebo	No of Participants (studies)
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean quality of sleep in the zaleplon group was **0.10 points**[2] **lower** (0.27 lower to 0.07 higher)	277 (1 study) [A]

* Critical Outcome, used to determine Quality of Evidence
[1] Study funded by Industry
[2] 7-point scale (1=excellent, 7=extremely poor)

Table S7 – Summary of Findings table for zaleplon 10 mg for the treatment of chronic insomnia

References: Hedner 2000(A); Walsh 2000(B)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 10 mg Zaleplon vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[1,3]	The mean sleep latency in the zaleplon group was **9.50 minutes lower** (18.80 to 0.19 lower)	94 (1 study)[B]
Sleep Latency (Subjective)	⊕⊕⊖⊖ **low**[2,3]	The mean sleep latency in the zaleplon group was **11.40 minutes lower** (27.36 lower to 4.56 higher)	92 (1 study)[B]
Total Sleep Time* (Subjective)	⊕⊕⊖⊖ **low**[3,4]	The mean total sleep time in the zaleplon group was **21.50 minutes higher** (5.60 lower to 48.6 higher)	93 (1 study) [B]
Wake After Sleep Onset (PSG)	⊕⊕⊕⊖ **moderate**[3]	The mean wake after sleep onset in the zaleplon group was **2.10 minutes lower** (10.23 lower to 6.03 higher)	92 (1 study)[B]
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean quality of sleep in the zaleplon group was **0.10 points**[5] **lower** (0.27 lower to 0.07 higher)	283 (1 study)[A]

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-18.8, -0.19) crosses Clinical Significance (10 min)
[2] 95% CI (-27.36, 4.56) crosses Clinical Significance (20 min)
[3] Study funded by Industry
[4] 95% CI (-5.60, 48.60) crosses Clinical Significance (30 min)
[5] 7-point scale (1=excellent, 7=extremely poor)

Zolpidem - Meta-Analyses and Summary of Findings Tables

Table S8 – Summary of Findings table for zolpidem 6.25 mg for the treatment of chronic insomnia

References: Walsh 2008

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 6.25 mg Zolpidem vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean sleep latency in the zolpidem group was **5.27 minutes lower** (11.47 lower to 0.93 higher)	199 (1 study)
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[1,3]	The mean wake after sleep onset in the zolpidem group was **13.03 minutes lower** (22.5 to 3.55 lower)	199 (1 study)
Sleep Efficiency (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep efficiency in the zolpidem group was **1.60 percent higher** (1.4 lower to 4.6 higher)	199 (1 study)

* Critical Outcome, used to determine Quality of Evidence
[1] Funding source not specified, author disclosures not specified.
[2] 95% CI (-11.47, 0.93) crosses Clinical Significance
[3] 95% CI (-22.5, -3.55) crosses Clinical Significance (20 min)

Figure S23 – Meta-analysis of data for PSG-determined sleep latency in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Herrmann, 1993	28	7	11	41.7	15	10	18.4%	-13.70 [-23.88, -3.52]
Randal 2012	14.2	12.2	44	29.76	26.9	47	20.5%	-15.56 [-24.05, -7.07]
Scharf, 1994	25.8	13.7	22	28.1	25.6	22	16.1%	-2.30 [-14.43, 9.83]
Uchimura 2012	14.3	22.6	70	37.5	37.8	71	18.3%	-23.20 [-33.46, -12.94]
Ware, 1997	37	6	34	42	5	35	26.8%	-5.00 [-7.61, -2.39]
Total (95% CI)			**181**			**185**	**100.0%**	**-11.65 [-19.15, -4.15]**

Heterogeneity: Tau² = 52.80; Chi² = 17.97, df = 4 (P = 0.001); I² = 78%
Test for overall effect: Z = 3.04 (P = 0.002)

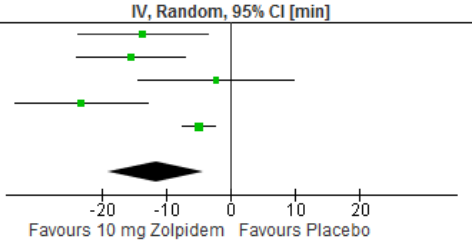


Figure S24 – Meta-analysis of data for subjectively-determined sleep latency in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Dorsey, 2004	29	2.64	68	37	3.16	73	17.2%	-8.00 [-8.96, -7.04]
Elie 1999	30.8	27.8	64	54.4	39.3	63	9.3%	-23.60 [-35.46, -11.74]
Erman, 2008	30.8	27.8	64	54.4	39.3	63	9.3%	-23.60 [-35.46, -11.74]
Herrmann, 1993	40.5	10	11	72.8	10	10	12.0%	-32.30 [-40.86, -23.74]
Jacobs, 2004	45	32.5	14	66.8	37.7	14	3.4%	-21.80 [-47.87, 4.27]
Perlis, 2004	38.4	33.1	95	55.1	52.3	97	9.0%	-16.70 [-29.05, -4.35]
Randal 2012	27.33	31.4	44	36.8	35	47	8.1%	-9.47 [-23.12, 4.18]
Scharf, 1994	38.4	22	22	56.6	39.5	23	5.6%	-18.20 [-36.78, 0.38]
Uchimura 2012	28	24.6	70	62	47.8	71	8.9%	-34.00 [-46.52, -21.48]
Walsh, 1998	48.1	3.1	91	64.7	4.6	97	17.2%	-16.60 [-17.72, -15.48]
Total (95% CI)			**543**			**558**	**100.0%**	**-19.55 [-24.90, -14.20]**

Heterogeneity: Tau² = 43.12; Chi² = 174.68, df = 9 (P < 0.00001); I² = 95%
Test for overall effect: Z = 7.16 (P < 0.00001)

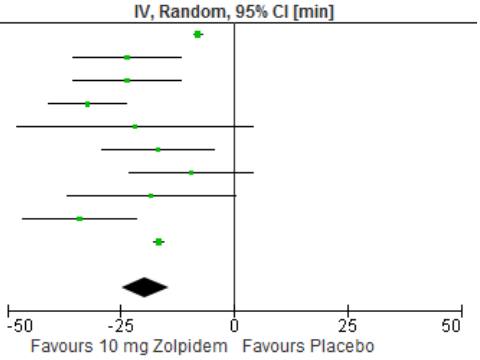


Figure S25 – Meta-analysis of data for PSG-determined total sleep time in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Herrmann, 1993	381.3	10	11	360.3	23	10	57.6%	21.00 [5.57, 36.43]
Randal 2012	411	32.7	44	371.36	65.1	47	42.4%	39.64 [18.67, 60.61]
Total (95% CI)			**55**			**57**	**100.0%**	**28.91 [10.85, 46.97]**

Heterogeneity: Tau² = 85.49; Chi² = 1.97, df = 1 (P = 0.16); I² = 49%
Test for overall effect: Z = 3.14 (P = 0.002)

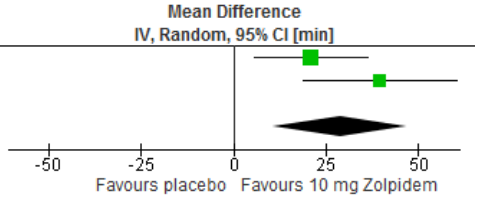


Figure S26 – Meta-analysis of data for subjectively-determined total sleep time in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Elie 1999	324.7	133	94	353	105	104	10.9%	-28.30 [-61.92, 5.32]
Erman, 2008	438.7	406.4	64	362	63.8	63	2.0%	76.70 [-24.10, 177.50]
Herrmann, 1993	372.7	12	11	327.4	22	10	19.4%	45.30 [29.93, 60.67]
Jacobs, 2004	343.4	76.8	14	296.8	99.6	14	4.2%	46.60 [-19.28, 112.48]
Perlis, 2004	417	64.4	95	359.8	77.1	97	17.0%	57.20 [37.12, 77.28]
Randal 2012	402	78	44	390	66	47	12.4%	12.00 [-17.79, 41.79]
Scharf, 1994	369	65	22	356	64	23	9.6%	13.00 [-24.71, 50.71]
Walsh, 1998	378.8	5.3	91	344.6	5.3	97	24.4%	34.20 [32.68, 35.72]
Total (95% CI)			**435**			**455**	**100.0%**	**30.04 [15.12, 44.96]**

Heterogeneity: Tau² = 236.58; Chi² = 24.49, df = 7 (P = 0.0009); I² = 71%
Test for overall effect: Z = 3.95 (P < 0.0001)

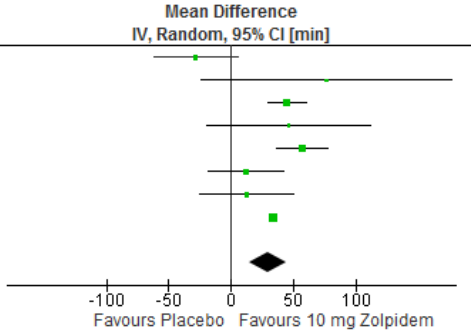


Figure S27 – Meta-analysis of data for PSG-determined wake after sleep onset in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Herrmann, 1993	34.7	7	11	60	12	10	78.2%	-25.30 [-33.81, -16.79]
Randal 2012	58.25	32.7	44	84.3	45.2	47	21.8%	-26.05 [-42.19, -9.91]
Total (95% CI)			**55**			**57**	**100.0%**	**-25.46 [-32.99, -17.94]**

Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.94); I² = 0%
Test for overall effect: Z = 6.63 (P < 0.00001)

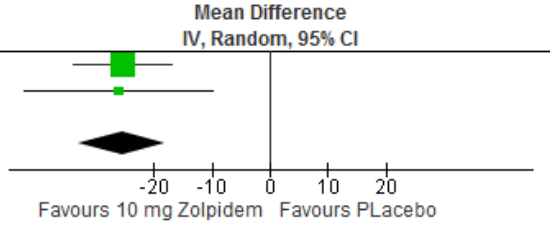


Figure S28 – Meta-analysis of data for subjectively-determined wake after sleep onset in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Dorsey, 2004	29	5.3	68	47	4.72	73	35.3%	-18.00 [-19.66, -16.34]
Erman, 2008	74.7	226.6	64	56.6	48.3	63	1.2%	18.10 [-38.68, 74.88]
Perlis, 2004	32.6	43.5	95	55.4	56.1	97	12.7%	-22.80 [-36.98, -8.62]
Randal 2012	52.3	58.3	44	59.62	52.4	47	6.2%	-7.32 [-30.15, 15.51]
Scharf, 1994	32.7	30.2	22	37.2	32.5	23	8.8%	-4.50 [-22.82, 13.82]
Walsh, 1998	39.5	3.6	91	49.8	4	97	35.8%	-10.30 [-11.39, -9.21]
Total (95% CI)			**384**			**400**	**100.0%**	**-13.57 [-19.84, -7.30]**

Heterogeneity: Tau² = 28.27; Chi² = 61.87, df = 5 (P < 0.00001); I² = 92%
Test for overall effect: Z = 4.24 (P < 0.0001)

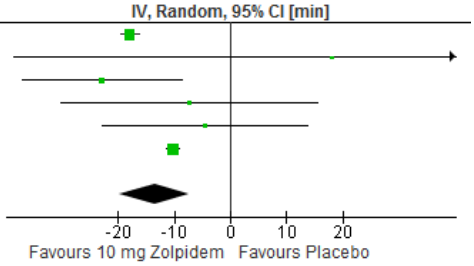


Figure S29 – Meta-analysis of data for subjectively-determined quality of sleep in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Erman, 2008	55.6	21.1	64	44.4	19.9	63	17.2%	0.54 [0.19, 0.90]
Randal 2012	1.4	0.7	44	1.38	0.6	47	16.9%	0.03 [-0.38, 0.44]
Scharf, 1994	1.5	0.7	22	1.4	0.5	23	15.7%	0.16 [-0.42, 0.75]
Staner, 2005	68.8	21.5	23	61.1	21.6	23	15.7%	0.35 [-0.23, 0.93]
Uchimura 2012	6.6	2	70	5.2	1.9	71	17.3%	0.71 [0.37, 1.05]
Walsh, 1998	1.55	0.05	91	1.44	0.06	97	17.2%	1.98 [1.63, 2.33]
Total (95% CI)			**314**			**324**	**100.0%**	**0.64 [0.03, 1.26]**

Heterogeneity: Tau² = 0.54; Chi² = 65.94, df = 5 (P < 0.00001); I² = 92%
Test for overall effect: Z = 2.05 (P = 0.04)

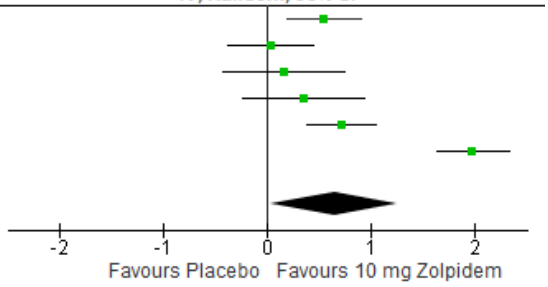


Figure S30 – Meta-analysis of data for PSG-determined sleep efficiency in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Herrmann, 1993	86.6	2	11	78.3	5	10	19.4%	8.30 [4.98, 11.62]
Randal 2012	85.71	6.5	44	78.88	10.4	47	17.6%	6.83 [3.29, 10.37]
Scharf, 1994	87.9	6.4	22	80.7	13.4	23	7.2%	7.20 [1.11, 13.29]
Ware, 1997	79	2	34	74	2	35	55.8%	5.00 [4.06, 5.94]
Total (95% CI)			**111**			**115**	**100.0%**	**6.12 [4.39, 7.85]**

Heterogeneity: Tau² = 1.17; Chi² = 4.59, df = 3 (P = 0.20); I² = 35%
Test for overall effect: Z = 6.93 (P < 0.00001)

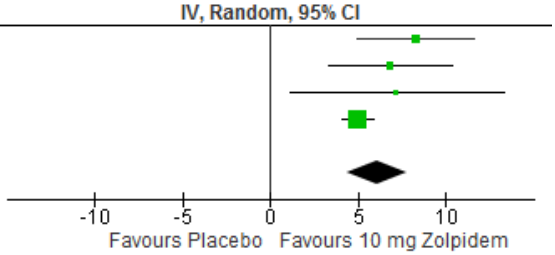


Figure S31 – Meta-analysis of data for PSG-determined number of awakenings in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Scharf, 1994	6.7	3.2	22	6.7	3.5	22	5.4%	0.00 [-1.98, 1.98]
Ware, 1997	6	1	34	7	1	35	94.6%	-1.00 [-1.47, -0.53]
Total (95% CI)			**56**			**57**	**100.0%**	**-0.95 [-1.41, -0.49]**

Heterogeneity: Tau² = 0.00; Chi² = 0.93, df = 1 (P = 0.34); I² = 0%
Test for overall effect: Z = 4.04 (P < 0.0001)

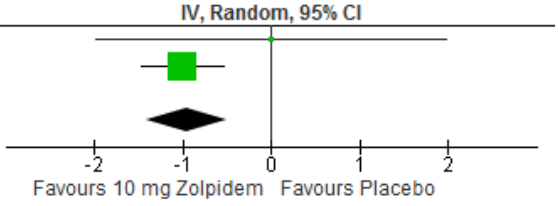


Figure S32 – Meta-analysis of data for subjectively-determined number of awakenings in response to zolpidem 10 mg

Study or Subgroup	10 mg Zolpidem Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Dorsey, 2004	1.4	0.12	68	1.8	0.12	73	34.1%	-0.40 [-0.44, -0.36]
Elie 1999	2.6	1.7	63	1.8	1	70	6.8%	0.80 [0.32, 1.28]
Herrmann, 1993	1.8	0.4	11	2.3	0.4	10	11.2%	-0.50 [-0.84, -0.16]
Perlis, 2004	1.03	0.92	95	1.64	1.33	97	12.1%	-0.61 [-0.93, -0.29]
Scharf, 1994	2.7	2.1	22	2.5	1	23	1.9%	0.20 [-0.77, 1.17]
Walsh, 1998	1.5	0.2	91	1.8	0.1	97	33.8%	-0.30 [-0.35, -0.25]
Total (95% CI)			**350**			**370**	**100.0%**	**-0.31 [-0.45, -0.17]**

Heterogeneity: Tau² = 0.01; Chi² = 37.05, df = 5 (P < 0.00001); I² = 87%
Test for overall effect: Z = 4.37 (P < 0.0001)

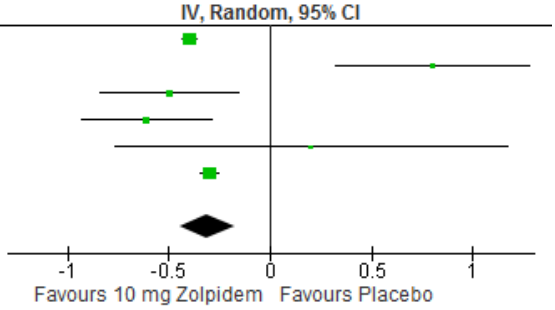


Figure S33 – Meta-analysis of data for the occurrence of amnesia in response to zolpidem 10 mg

Study or Subgroup	Zolpidem 10mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Elie 1999	6	122	3	126	40.0%	0.03 [-0.02, 0.07]
Scharf 1994	5	122	0	126	60.0%	0.04 [0.00, 0.08]
Total (95% CI)		**244**		**252**	**100.0%**	**0.03 [0.01, 0.06]**
Total events	11		3			

Heterogeneity: Tau² = 0.00; Chi² = 0.27, df = 1 (P = 0.60); I² = 0%
Test for overall effect: Z = 2.31 (P = 0.02)

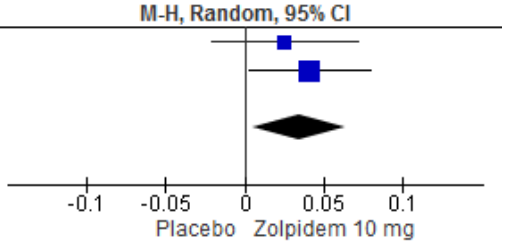


Figure S34 – Meta-analysis of data for the occurrence of dizziness in response to zolpidem 10 mg

Study or Subgroup	Zolpidem 10mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Dorsey 2004	6	68	0	73	27.2%	0.09 [0.02, 0.16]
Erman JCSM 2008	6	64	3	63	17.7%	0.05 [-0.04, 0.13]
Scharf 1994	3	26	0	24	7.3%	0.12 [-0.02, 0.25]
Uchimura 2012	3	70	0	71	47.8%	0.04 [-0.01, 0.10]
Total (95% CI)		**228**		**231**	**100.0%**	**0.06 [0.02, 0.10]**
Total events	18		3			

Heterogeneity: Tau² = 0.00; Chi² = 1.75, df = 3 (P = 0.62); I² = 0%
Test for overall effect: Z = 3.21 (P = 0.001)

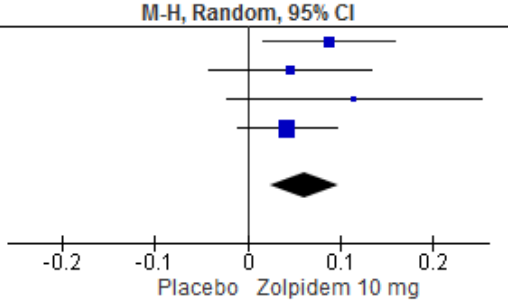


Figure S35 – Meta-analysis of data for the occurrence of headache in response to zolpidem 10 mg

Study or Subgroup	Zolpidem 10mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Dorsey 2004	36	68	24	73	26.4%	0.20 [0.04, 0.36]
Erman JCSM 2008	6	64	6	63	38.6%	-0.00 [-0.10, 0.10]
Walsh 1998	22	91	18	97	35.0%	0.06 [-0.06, 0.17]
Total (95% CI)		**223**		**233**	**100.0%**	**0.07 [-0.04, 0.19]**
Total events	64		48			

Heterogeneity: Tau² = 0.01; Chi² = 5.00, df = 2 (P = 0.08); I² = 60%
Test for overall effect: Z = 1.25 (P = 0.21)

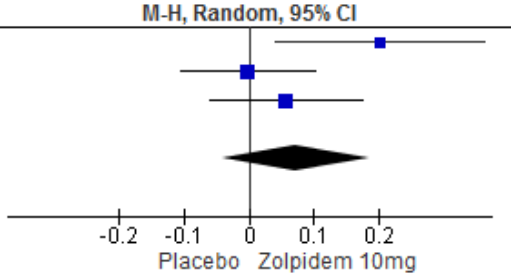


Figure S36 – Meta-analysis of data for the occurrence of nausea in response to zolpidem 10 mg

Study or Subgroup	Zolpidem 10mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Erman JCSM 2008	4	64	2	63	68.7%	0.03 [-0.04, 0.10]
Scharf 1994	1	26	1	24	31.3%	-0.00 [-0.11, 0.11]
Total (95% CI)		**90**		**87**	**100.0%**	**0.02 [-0.04, 0.08]**
Total events	5		3			

Heterogeneity: Tau² = 0.00; Chi² = 0.26, df = 1 (P = 0.61); I² = 0%
Test for overall effect: Z = 0.65 (P = 0.52)

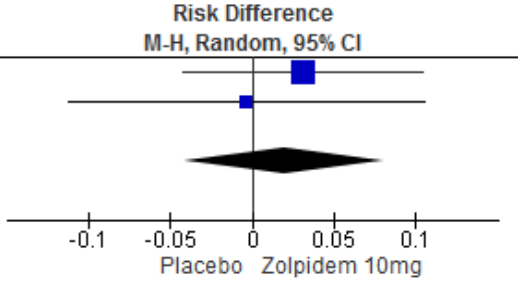


Figure S37 – Meta-analysis of data for the occurrence of somnolence in response to zolpidem 10 mg

Study or Subgroup	Zolpidem 10mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Dorsey 2004	7	68	1	73	14.0%	0.09 [0.01, 0.17]
Elie 1999	6	122	3	126	37.9%	0.03 [-0.02, 0.07]
Erman JCSM 2008	6	64	2	63	11.9%	0.06 [-0.02, 0.15]
Scharf 1994	3	26	0	24	4.3%	0.12 [-0.02, 0.25]
Uchimura 2012	3	70	2	71	22.2%	0.01 [-0.05, 0.08]
Walsh 1998	14	91	8	97	9.7%	0.07 [-0.02, 0.16]
Total (95% CI)		**441**		**454**	**100.0%**	**0.04 [0.02, 0.07]**
Total events	39		16			

Heterogeneity: Tau² = 0.00; Chi² = 4.74, df = 5 (P = 0.45); I² = 0%
Test for overall effect: Z = 3.04 (P = 0.002)

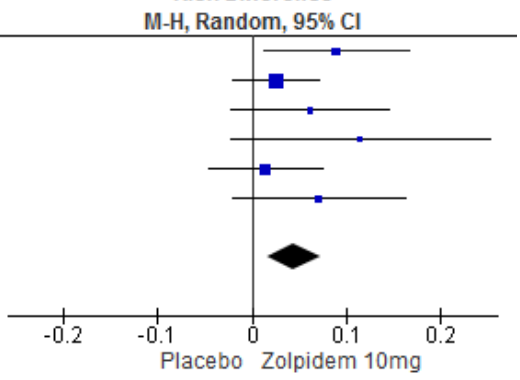


Figure S38 – Meta-analysis of data for the occurrence of taste perversion in response to zolpidem 10 mg

Study or Subgroup	Zolpidem 10mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Elie 1999	3	122	1	126	60.5%	0.02 [-0.01, 0.05]
Uchimura 2012	1	70	1	71	39.5%	0.00 [-0.04, 0.04]
Total (95% CI)		**192**		**197**	**100.0%**	**0.01 [-0.01, 0.03]**
Total events	4		2			

Heterogeneity: Tau² = 0.00; Chi² = 0.41, df = 1 (P = 0.52); I² = 0%
Test for overall effect: Z = 0.81 (P = 0.42)

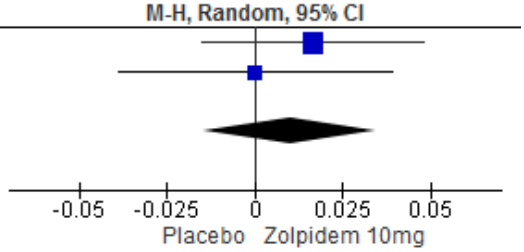


Table S9 – Summary of Findings table for zolpidem 10 mg for the treatment of chronic insomnia

References: Dorsey 2004(A); Elie 1999(B); Erman 2008(C); Herrmann 1993(D); Jacobs 2004(E); Perlis 2004(F); Randal 2012(G); Scharf 1994(H); Staner 2005(I); Uchimura 2012(J); Walsh 1998(K); Ware 1997(L)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 10 mg Zolpidem vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊖⊖⊖ **very low**[1,2,3]	The mean sleep latency in the zolpidem groups was **11.65 minutes lower** (19.15 to 4.15 lower)	366 (5 studies)[D,G,H,J,L]
Sleep Latency (Subjective)	⊕⊖⊖⊖ **very low**[3,4,5]	The mean sleep latency in the zolpidem groups was **19.55 minutes lower** (24.90 to 14.20 lower)	1101 (10 studies)[A,B,C,D,E,FG,H,J,K]
Total Sleep Time* (PSG)	⊕⊕⊖⊖ **low**[3,12]	The mean total sleep time in the zolpidem groups was **28.91 minutes higher** (10.85 to 46.97 higher)	112 (2 studies)[D,G]
Total Sleep Time* (Subjective)	⊕⊕⊖⊖ **low**[3,7]	The mean total sleep time in the zolpidem groups was **30.04 minutes higher** (15.12 to 44.96 higher)	890 (8 studies)[B,C,D,E,F,G,H,K]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[3,13]	The mean wake after sleep onset in the zolpidem groups was **25.46 minutes lower** (32.99 to 17.94 lower)	112 (2 studies)[D,G]
Wake After Sleep Onset (Subjective)	⊕⊕⊖⊖ **low**[3,6]	The mean wake after sleep onset in the zolpidem groups was **13.57 minutes lower** (19.84 to 7.30 lower)	784 (6 studies)[A,C,F,G,H,K]
Quality of Sleep* (Subjective)	⊕⊖⊖⊖ **very low**[3,10,11]	The mean quality of sleep in the zolpidem groups was **0.64 standard deviations higher** (0.03 to 1.26 higher)	638 (6 studies)[C,G,H,I,J,K]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[3,9]	The mean sleep efficiency in the zolpidem groups was **6.12 percent higher** (4.39 to 7.85 higher)	226 (4 studies)[D,G,H,L]
Number of Awakenings (PSG)	⊕⊕⊕⊖ **moderate**[3]	The mean number of awakenings in the zolpidem groups was **0.95 awakenings lower** (1.41 to 0.49 lower)	113 (2 studies)[H,L]
Number of Awakenings (Subjective)	⊕⊕⊖⊖ **low**[3,8]	The mean number of awakenings in the zolpidem groups was **0.31 awakenings lower** (0.45 to 0.17 lower)	720 (6 studies)[A,B,D,F,H,K]

* Critical Outcome, used to determine Quality of Evidence
[1] Heterogeneity (I^2 = 78%) greater than allowance (75%)
[2] 95% CI (-19.15, -4.15) crosses Clinical Significance (10 min)
[3] Studies funded by industry
[4] Heterogeneity (I^2 = 95%) greater than allowance (75%)
[5] 95% CI (-24.90, -14.20) crosses Clinical Significance (20 min)
[6] Heterogeneity (I^2 = 92%) greater than allowance (75%)
[7] 95% CI (15.12, 44.96) crosses Clinical Significance (30 min)
[8] Heterogeneity (I^2 = 87%) greater than allowance (75%)
[9] 95% CI (4.39, 7.85) crosses Clinical Significance (5%)
[10] Heterogeneity (I^2 = 92%) greater than allowance (75%)
[11] 95% CI (0.3, 1.26) crosses Clinical Significance (SMD 0.5)
[12] 95% CI (10.85, 46.97) crosses Clinical Significance (20 min)
[13] 95% CI (-32.99, -17.4) crosses Clinical Significance (20 min)

Table S10 – Summary of Findings table for zolpidem 12.5 mg for the treatment of chronic insomnia

References: Roth 2006

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 12.5 Zolpidem vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean sleep latency in the zolpidem group was **8.19 minutes lower** (15.22 to 1.15 lower)	212 (1 study)
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[1,3]	The mean wake after sleep onset in the zolpidem group was **19.99 minutes lower** (27.33 to 12.64 lower)	212 (1 study)
Sleep Efficiency (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep efficiency in the zolpidem group was **3.9 percent higher** (1.38 to 6.41 higher)	212 (1 study)

* Critical Outcome, used to determine Quality of Evidence
[1] Funding source not specified, author disclosures not specified.
[2] 95% CI (-15.22, 1.15) crosses Clinical Significance (10 min)
[3] 95% CI (-27.33, -12.64) crosses Clinical Significance (20 min)

Triazolam - Summary of Findings Table

Table S11 – Summary of Findings table for triazolam 0.25 mg for the treatment of chronic insomnia

References: Roehrs 2001

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 0.25 mg Triazolam vs Placebo	No of Participants (studies)
Sleep Latency* (Subjective)	⊕⊕⊕⊕ **high**	The mean sleep latency in the triazolam group was **9.20 minutes lower** (22.3 lower to 3.9 higher)	64 (1 study)
Total Sleep Time (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean total sleep time in the triazolam group was **25.20 minutes higher** (9.12 lower to 59.52 higher)	64 (1 study)
Quality of Sleep* (Subjective)	⊕⊕⊕⊕ **high**	The mean quality of sleep in the triazolam group was **0.37 points**[3] **lower** (0.66 to 0.07 lower)	64 (1 study)
Number of Awakenings (Subjective)	⊕⊕⊖⊖ **low**[2]	The mean number of awakenings in the triazolam group was **0.37 awakenings lower** (1.7 lower to 0.96 higher)	64 (1 study)

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-9.12, 59.52) crosses Clinical Significance (30 min)
[2] 95% CI (-1.7, 0.96) crosses Clinical Significance (0.5 awakenings)
[3] 4-point scale (1=good, 4=poor)

Temazepam - Meta-Analyses and Summary of Findings Tables

Figure S39 – Meta-analysis of data for subjectively-determined sleep latency in response to temazepam 15 mg

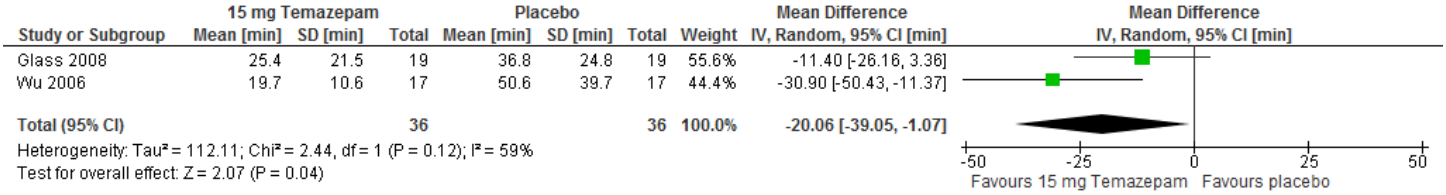

15 mg Temazepam
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Glass 2008 25.4 21.5 19 36.8 24.8 19 55.6% -11.40 [-26.16, 3.36]
Wu 2006 19.7 10.6 17 50.6 39.7 17 44.4% -30.90 [-50.43, -11.37]
Total (95% CI) 36 36 100.0% -20.06 [-39.05, -1.07]
Heterogeneity: Tau² = 112.11; Chi² = 2.44, df = 1 (P = 0.12); I² = 59%
Test for overall effect: Z = 2.07 (P = 0.04)
-50
-25
0
25
50
Favours 15 mg Temazepam
Favours placebo

Figure S40 – Meta-analysis of data for subjectively-determined total sleep time in response to temazepam 15 mg

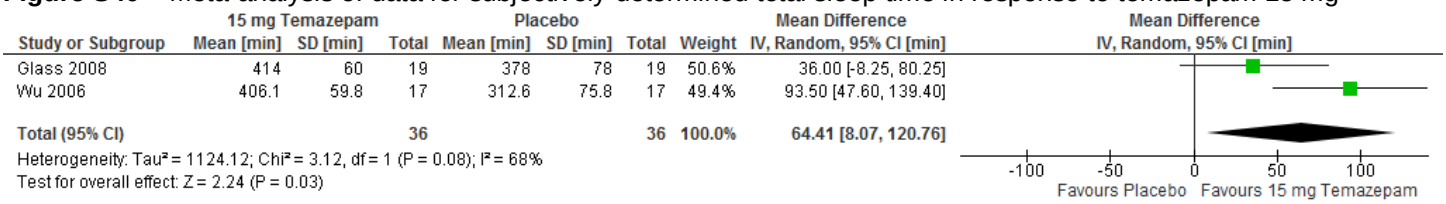

15 mg Temazepam
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Glass 2008 414 60 19 378 78 19 50.6% 36.00 [-8.25, 80.25]
Wu 2006 406.1 59.8 17 312.6 75.8 17 49.4% 93.50 [47.60, 139.40]
Total (95% CI) 36 36 100.0% 64.41 [8.07, 120.76]
Heterogeneity: Tau² = 1124.12; Chi² = 3.12, df = 1 (P = 0.08); I² = 68%
Test for overall effect: Z = 2.24 (P = 0.03)
-100
-50
0
50
100
Favours Placebo
Favours 15 mg Temazepam

Figure S41 – Meta-analysis of data for subjectively-determined quality of sleep in response to temazepam 15 mg

Study or Subgroup	15 mg Temazepam Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Glass 2008	3.3	0.86	19	2.9	0.77	19	47.9%	0.48 [-0.17, 1.13]
Hindmarch 1979	4.67	0.94	20	4.62	1.43	20	52.1%	0.04 [-0.58, 0.66]
Total (95% CI)			**39**			**39**	**100.0%**	**0.25 [-0.20, 0.70]**

Heterogeneity: Tau² = 0.00; Chi² = 0.92, df = 1 (P = 0.34); I² = 0%
Test for overall effect: Z = 1.10 (P = 0.27)

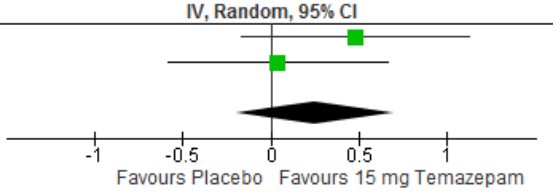


Table S12 – Summary of Findings table for temazepam 15 mg for the treatment of chronic insomnia

Reference: Glass 2008(A); Hindmarch 1979(B); Wu 2006 (C)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 15 mg Temazepam vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊕⊕ **high**	The mean sleep latency in the temazepam group was **37.1 minutes lower** (52.8 to 21.31 lower)	34 (1 study)[C]
Sleep Latency (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean sleep latency in the temazepam group was **20.06 minutes lower** (39.05 to 1.07 lower)	72 (2 studies)[A,C]
Total Sleep Time* (PSG)	⊕⊕⊕⊕ **high**	The mean total sleep time in the temazepam group was **99.1 minutes higher** (63.4 to 134.7 lower)	34 (1 study)[C]
Total Sleep Time (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean total sleep time in the temazepam groups was **64.41 minutes higher** (8.07 to 120.76 higher)	72 (2 studies)[A,C]
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean quality of sleep in the temazepam group was **0.25 standard deviations higher** (0.2 lower to 0.7 higher)	39 (2 studies)[A,B]
Sleep Efficiency (PSG)	⊕⊕⊕⊖ **moderate**[5]	The mean sleep efficiency in the temazepam group was **13.3 percent higher** (3.9 to 22.6 higher)	34 (1 study)[C]
Sleep Efficiency (Subjective)	⊕⊕⊕⊖ **moderate**[6]	The mean sleep efficiency in the temazepam group was **14.1 percent higher** (5.8 to 22.3 higher)	34 (1 study)[C]
Number of Awakenings (Subjective)	⊕⊕⊕⊖ **moderate**[4]	The mean number of awakenings in the temazepam group was **0.5 awakenings lower** (1.29 lower to 0.29 higher)	38 (1 study)[A]

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-0.2, 0.7) crosses Clinical Significance (0.5 SMD)
[2] 95% CI (-39.05, -1.07) crosses Clinical Significance (20 min)
[3] 95% CI (8.07,120.76) crosses Clinical Significance (30 min)
[4] 95% CI (-1.29, 0.29) crosses Clinical Significance (0.5 awakenings)
[5] 95% CI (3.9, 22.6) crosses Clinical Significance (5%)
[6] 95% CI (5.8, 22.3) crosses Clinical Significance (10%)

Table S13 – Summary of Findings table for temazepam 30 mg for the treatment of chronic insomnia

References: Hindmarch 1979

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 30 mg Temazepam vs Placebo	No of Participants (studies)
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean quality of sleep in the temazepam group was **0.69 cm² higher** (0.28 lower to 1.66 higher)	40 (1 study)

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-0.28, 1.66) crosses Clinical Significance (1.0 cm)
[2] 10 cm line analogue rating scale

Ramelteon - Meta-Analyses and Summary of Findings Table

Figure S42 – Meta-analysis of data for PSG-determined sleep latency in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Kohsaka, 2011	22.12	18.16	61	35.27	40.68	61	6.9%	-13.15 [-24.33, -1.97]
Roth, 2007	30.8	2.52	100	38.4	2.49	100	46.5%	-7.60 [-8.29, -6.91]
Zammit, 2007	31.5	2.91	139	42.5	2.97	131	46.5%	-11.00 [-11.70, -10.30]
Total (95% CI)			**300**			**292**	**100.0%**	**-9.57 [-12.75, -6.38]**

Heterogeneity: Tau² = 5.55; Chi² = 46.01, df = 2 (P < 0.00001); I² = 96%
Test for overall effect: Z = 5.89 (P < 0.00001)

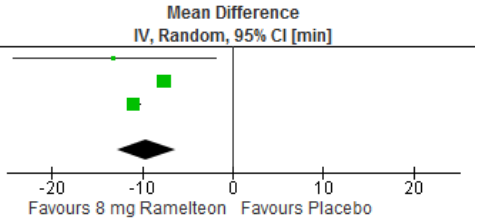


Figure S43 – Meta-analysis of data for subjectively-determined sleep latency in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Kohsaka, 2011	37.8	25.51	61	47.07	41.27	61	21.0%	-9.27 [-21.45, 2.91]
Roth, 2007	50.9	4.46	100	58.2	4.53	100	39.4%	-7.30 [-8.55, -6.05]
Zammit, 2007	44.8	3.6	139	61.5	3.7	131	39.6%	-16.70 [-17.57, -15.83]
Total (95% CI)			**300**			**292**	**100.0%**	**-11.44 [-19.56, -3.31]**

Heterogeneity: Tau² = 43.20; Chi² = 147.30, df = 2 (P < 0.00001); I² = 99%
Test for overall effect: Z = 2.76 (P = 0.006)

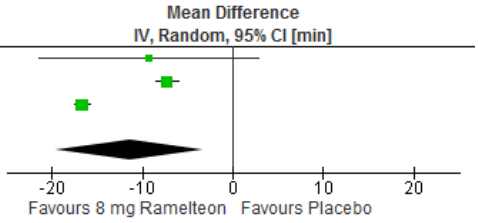


Figure S44 – Meta-analysis of data for PSG-determined total sleep time in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Kohsaka, 2011	407.2	45.33	61	394.7	52.35	61	7.0%	12.50 [-4.88, 29.88]
Mayer, 2009	381.39	3.233	159	380.11	3.255	176	31.2%	1.28 [0.58, 1.98]
Roth, 2007	362	5.03	100	350.4	5.04	100	30.7%	11.60 [10.20, 13.00]
Zammit, 2007	391.5	4.04	139	385.9	4.12	131	31.1%	5.60 [4.63, 6.57]
Total (95% CI)			**459**			**468**	**100.0%**	**6.58 [1.36, 11.80]**

Heterogeneity: Tau² = 22.59; Chi² = 183.95, df = 3 (P < 0.00001); I² = 98%
Test for overall effect: Z = 2.47 (P = 0.01)

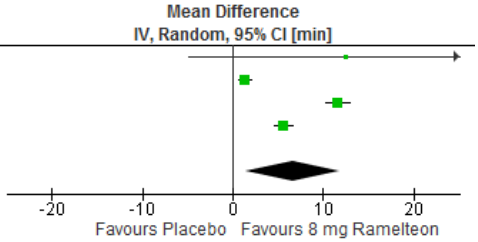


Figure S45 – Meta-analysis of data for subjectively-determined total sleep time in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Kohsaka, 2011	377.01	70.92	61	371.68	61.06	61	15.1%	5.33 [-18.15, 28.81]
Mayer, 2009	345.39	4.336	159	349.49	4.346	176	28.3%	-4.10 [-5.03, -3.17]
Roth, 2007	337	6.62	100	333.9	6.69	100	28.2%	3.10 [1.26, 4.94]
Zammit, 2007	365.4	5.6	139	347.1	5.7	131	28.3%	18.30 [16.95, 19.65]
Total (95% CI)			**459**			**468**	**100.0%**	**5.70 [-7.65, 19.04]**

Heterogeneity: Tau² = 163.34; Chi² = 717.67, df = 3 (P < 0.00001); I² = 100%
Test for overall effect: Z = 0.84 (P = 0.40)

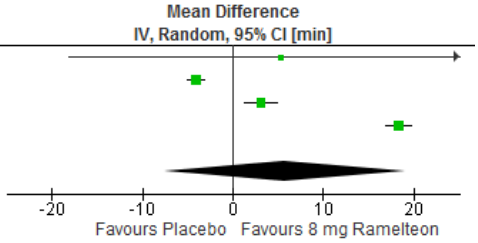


Figure S46 – Meta-analysis of data for PSG-determined wake after sleep onset in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Kohsaka, 2011	49.78	39.1	61	46.42	33.49	61	0.3%	3.36 [-9.56, 16.28]
Zammit, 2007	59.9	3.04	139	56.4	3.11	131	99.7%	3.50 [2.77, 4.23]
Total (95% CI)			**200**			**192**	**100.0%**	**3.50 [2.77, 4.23]**

Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.98); I² = 0%
Test for overall effect: Z = 9.36 (P < 0.00001)

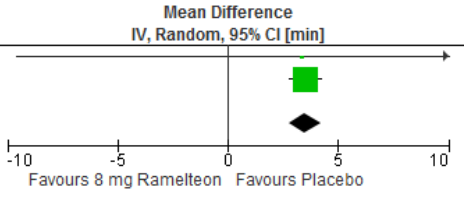


Figure S47 – Meta-analysis of data for subjectively-determined wake after sleep onset in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Mayer, 2009	90.89	3.796	159	79.54	3.822	176	50.1%	11.35 [10.53, 12.17]
Zammit, 2007	70.3	4.7	139	71.2	4.8	131	49.9%	-0.90 [-2.03, 0.23]
Total (95% CI)			**298**			**307**	**100.0%**	**5.23 [-6.77, 17.24]**

Heterogeneity: Tau² = 74.78; Chi² = 295.15, df = 1 (P < 0.00001); I² = 100%
Test for overall effect: Z = 0.85 (P = 0.39)

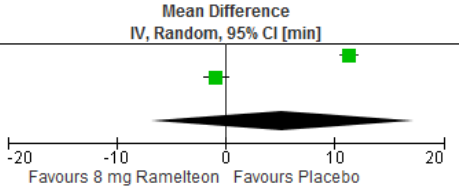


Figure S48 – Meta-analysis of data for PSG-determined quality of sleep in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Mayer, 2009	4.01	0.065	159	4.01	0.065	176	43.0%	0.00 [-0.01, 0.01]
Roth, 2007	3.8	0.1	100	3.8	1	100	14.1%	0.00 [-0.20, 0.20]
Zammit, 2007	3.6	0.06	139	3.7	0.06	131	42.9%	-0.10 [-0.11, -0.09]
Total (95% CI)			**398**			**407**	**100.0%**	**-0.04 [-0.13, 0.05]**

Heterogeneity: Tau² = 0.00; Chi² = 96.43, df = 2 (P < 0.00001); I² = 98%
Test for overall effect: Z = 0.93 (P = 0.35)

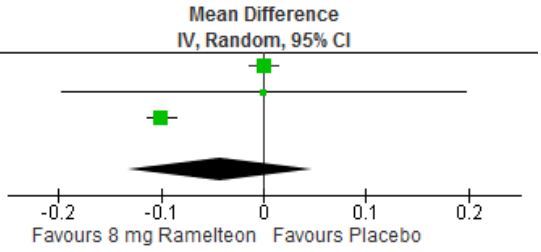


Figure S49 – Meta-analysis of data for PSG-determined sleep efficiency in response to ramelteon 8 mg

Study or Subgroup	8 mg Ramelteon Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Kohsaka, 2011	84.83	9.44	61	82.26	10.9	61	5.9%	2.57 [-1.05, 6.19]
Roth, 2007	75.5	1.05	100	73.1	1.05	100	46.5%	2.40 [2.11, 2.69]
Zammit, 2007	81.8	0.84	139	80.4	0.86	131	47.6%	1.40 [1.20, 1.60]
Total (95% CI)			**300**			**292**	**100.0%**	**1.93 [1.00, 2.87]**

Heterogeneity: Tau² = 0.47; Chi² = 30.72, df = 2 (P < 0.00001); I² = 93%
Test for overall effect: Z = 4.05 (P < 0.0001)

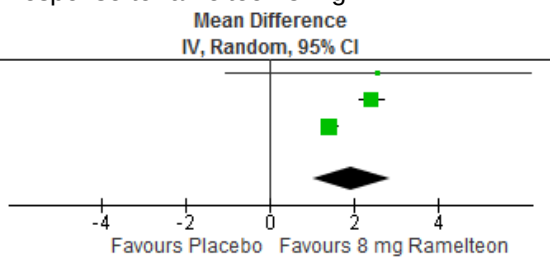


Figure S50 – Meta-analysis of data for the occurrence of headache in response to ramelteon 8 mg

Study or Subgroup	Ramelteon 8 mg Events	Total	placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Mayer 2009	18	228	18	223	33.8%	-0.00 [-0.05, 0.05]
Roth 2007	3	100	1	100	56.5%	0.02 [-0.02, 0.06]
Zammit 2007	27	139	24	131	9.7%	0.01 [-0.08, 0.10]
Total (95% CI)		**467**		**454**	**100.0%**	**0.01 [-0.02, 0.04]**
Total events	48		43			

Heterogeneity: Tau² = 0.00; Chi² = 0.57, df = 2 (P = 0.75); I² = 0%
Test for overall effect: Z = 0.79 (P = 0.43)

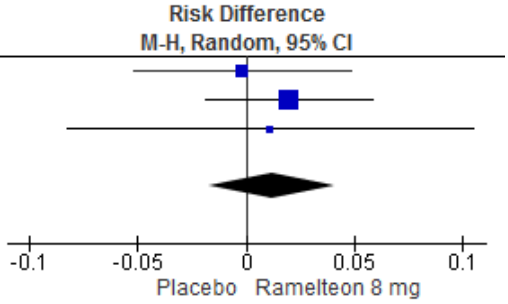


Figure S51 – Meta-analysis of data for the occurrence of upper respiratory tract infection in response to ramelteon 8 mg

Study or Subgroup	Ramelteon 8 mg Events	Total	placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Mayer 2009	18	228	18	223	44.5%	-0.00 [-0.05, 0.05]
Zammit 2007	6	139	4	131	55.5%	0.01 [-0.03, 0.06]
Total (95% CI)		**367**		**354**	**100.0%**	**0.01 [-0.03, 0.04]**
Total events	24		22			

Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 1 (P = 0.65); I² = 0%
Test for overall effect: Z = 0.37 (P = 0.72)

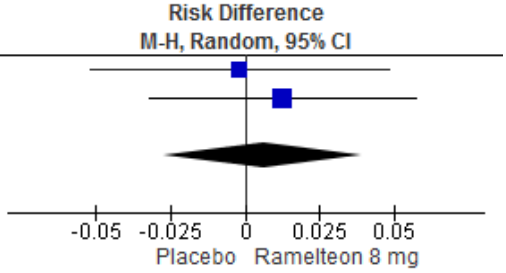


Table S14 – Summary of Findings table for ramelteon 8 mg for the treatment of chronic insomnia

References: Kohsaka 2011 (A); Mayer 2009(B); Roth 2007(C); Zammit 2007(D)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 8 mg Ramelteon vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊖⊖⊖ **very low**[1,2,3]	The mean sleep latency in the ramelteon groups was **9.57 minutes lower** (12.75 to 6.38 lower)	592 (3 studies) [A,C,D]
Sleep Latency (Subjective)	⊕⊕⊖⊖ **low**[3,7,8]	The mean sleep latency in the ramelteon groups was **11.44 minutes lower** (19.56 to 3.31 lower)	592 (3 studies) [A,C,D]
Total Sleep Time (PSG)	⊕⊕⊖⊖ **low**[1,3]	The mean total sleep time in the ramelteon groups was **6.58 minutes higher** (1.36 to 11.80 higher)	927 (4 studies) [A,B,C,D]
Total Sleep Time (Subjective)	⊕⊕⊖⊖ **low**[3,6]	The mean total sleep time in the ramelteon groups was **5.70 minutes higher** (7.65 lower to 19.04 higher)	927 (4 studies) [A,B,C,D]
Wake After Sleep Onset (PSG)	⊕⊕⊕⊖ **moderate**[3]	The mean wake after sleep onset in the ramelteon groups was **3.50 minutes higher** (2.77 to 4.23 higher)	392 (2 study) [A,D]
Wake After Sleep Onset (Subjective)	⊕⊕⊖⊖ **low**[3,6]	The mean wake after sleep onset in the ramelteon groups was **5.23 minutes higher** (6.77 lower to 17.24 higher)	605 (2 studies) [B,D]
Quality of Sleep* (Subjective)	⊕⊕⊖⊖ **low**[1,3]	The mean quality of sleep in the ramelteon groups was **0.04 points lower**[5] (0.13 lower to 0.05 higher)	805 (3 studies) [B,C,D]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[3,4]	The mean sleep efficiency in the ramelteon groups was **1.93 percent higher** (1.00 to 2.87 higher)	592 (3 studies) [A,C,D]
Number of Awakenings (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean number of awakenings in the ramelteon group was **0.12 awakenings higher** (0.08 to 0.15 higher)	335 (1 study) [B]

* Critical Outcome, used to determine Quality of Evidence
[1] Heterogeneity (I^2 = 98%) is greater than allowance (75%)
[2] 95% CI (-12.75, -6.38) crosses Clinical Significance (10 min)
[3] All studies funded by industry
[4] Heterogeneity (I^2 = 93%) greater than allowance (75%)
[5] 7-point Likert scale (1=excellent, 7=very poor)
[6] Heterogeneity (I^2 =100%) greater than allowance (75%)
[7] Heterogeneity (I^2 =99%) greater than allowance (75%)
[8] 95% CI (-21.45, 2.90) crossses Clinical Significance (20 min)

Doxepin - Meta-Analyses and Summary of Findings Tables

Figure S4952 – Meta-analysis of data for PSG-determined sleep latency in response to doxepin 3 mg

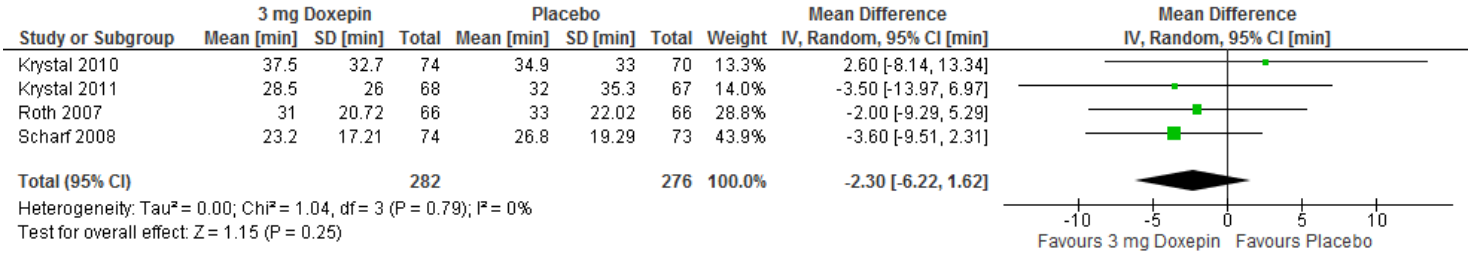

3 mg Doxepin
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Krystal 2010 37.5 32.7 74 34.9 33 70 13.3% 2.60 [-8.14, 13.34]
Krystal 2011 28.5 26 68 32 35.3 67 14.0% -3.50 [-13.97, 6.97]
Roth 2007 31 20.72 66 33 22.02 66 28.8% -2.00 [-9.29, 5.29]
Scharf 2008 23.2 17.21 74 26.8 19.29 73 43.9% -3.60 [-9.51, 2.31]
Total (95% CI) 282 276 100.0% -2.30 [-6.22, 1.62]
Heterogeneity: Tau² = 0.00; Chi² = 1.04, df = 3 (P = 0.79); I² = 0%
Test for overall effect: Z = 1.15 (P = 0.25)
-10 -5 0 5 10
Favours 3 mg Doxepin Favours Placebo

Figure S53 – Meta-analysis of data for subjectively-determined sleep latency in response to doxepin 3 mg

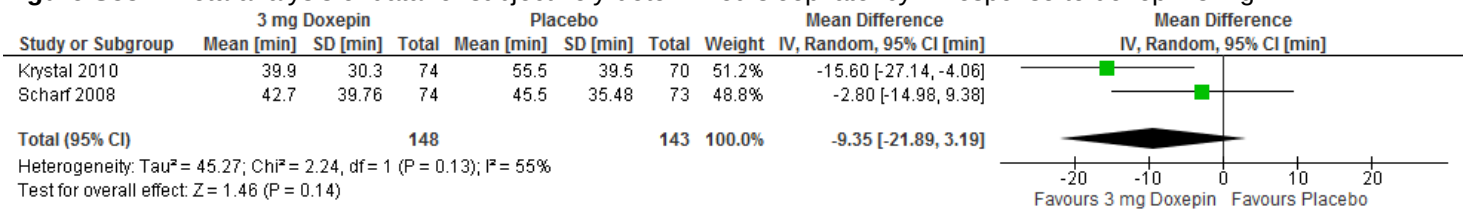

3 mg Doxepin
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Krystal 2010 39.9 30.3 74 55.5 39.5 70 51.2% -15.60 [-27.14, -4.06]
Scharf 2008 42.7 39.76 74 45.5 35.48 73 48.8% -2.80 [-14.98, 9.38]
Total (95% CI) 148 143 100.0% -9.35 [-21.89, 3.19]
Heterogeneity: Tau² = 45.27; Chi² = 2.24, df = 1 (P = 0.13); I² = 55%
Test for overall effect: Z = 1.46 (P = 0.14)
-20 -10 0 10 20
Favours 3 mg Doxepin Favours Placebo

Figure S54 – Meta-analysis of data for PSG-determined total sleep time in response to doxepin 3 mg

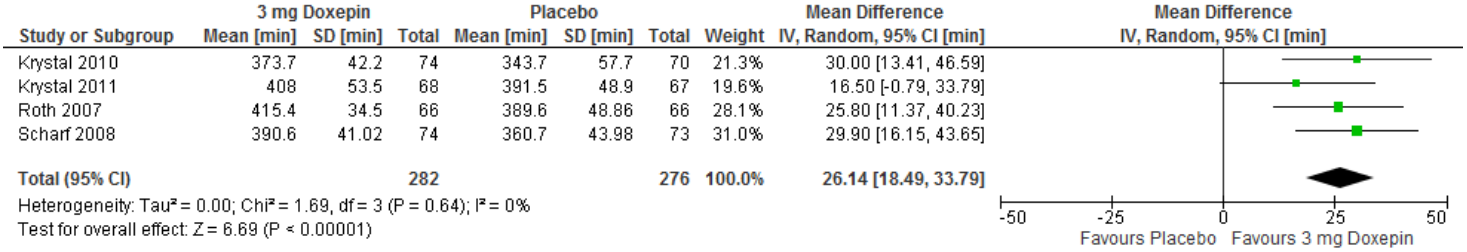

3 mg Doxepin
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Krystal 2010 373.7 42.2 74 343.7 57.7 70 21.3% 30.00 [13.41, 46.59]
Krystal 2011 408 53.5 68 391.5 48.9 67 19.6% 16.50 [-0.79, 33.79]
Roth 2007 415.4 34.5 66 389.6 48.86 66 28.1% 25.80 [11.37, 40.23]
Scharf 2008 390.6 41.02 74 360.7 43.98 73 31.0% 29.90 [16.15, 43.65]
Total (95% CI) 282 276 100.0% 26.14 [18.49, 33.79]
Heterogeneity: Tau² = 0.00; Chi² = 1.69, df = 3 (P = 0.64); I² = 0%
Test for overall effect: Z = 6.69 (P < 0.00001)
-50 -25 0 25 50
Favours Placebo Favours 3 mg Doxepin

Figure S55 – Meta-analysis of data for subjectively-determined total sleep time in response to doxepin 3 mg

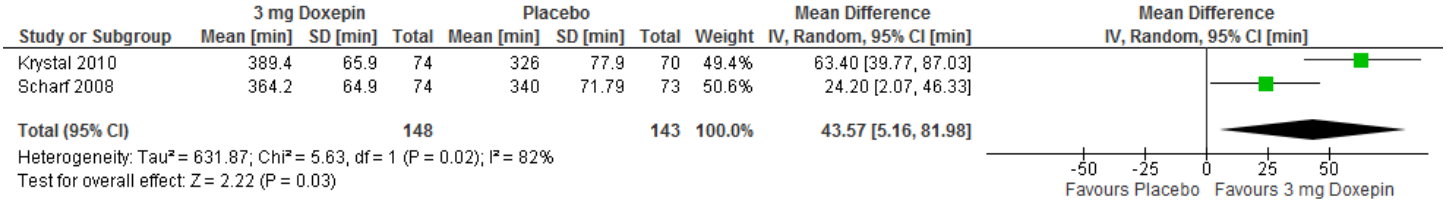

3 mg Doxepin
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Krystal 2010 389.4 65.9 74 326 77.9 70 49.4% 63.40 [39.77, 87.03]
Scharf 2008 364.2 64.9 74 340 71.79 73 50.6% 24.20 [2.07, 46.33]
Total (95% CI) 148 143 100.0% 43.57 [5.16, 81.98]
Heterogeneity: Tau² = 631.87; Chi² = 5.63, df = 1 (P = 0.02); I² = 82%
Test for overall effect: Z = 2.22 (P = 0.03)
-50 -25 0 25 50
Favours Placebo Favours 3 mg Doxepin

Figure S56 – Meta-analysis of data for PSG-determined wake after sleep onset in response to doxepin 3 mg

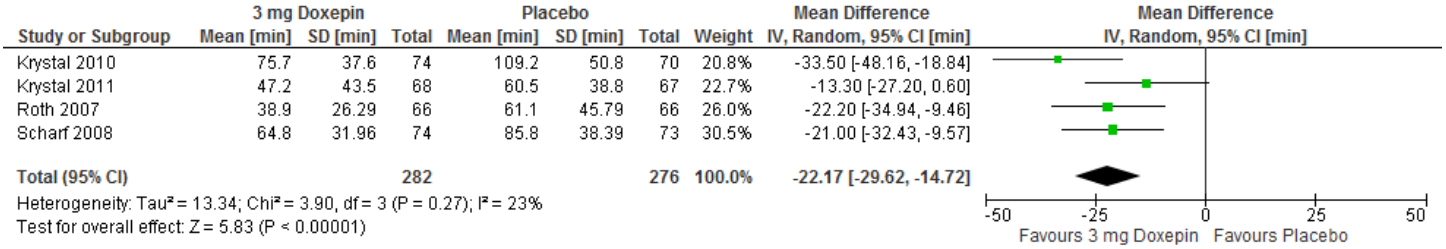

3 mg Doxepin
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Krystal 2010 75.7 37.6 74 109.2 50.8 70 20.8% -33.50 [-48.16, -18.84]
Krystal 2011 47.2 43.5 68 60.5 38.8 67 22.7% -13.30 [-27.20, 0.60]
Roth 2007 38.9 26.29 66 61.1 45.79 66 26.0% -22.20 [-34.94, -9.46]
Scharf 2008 64.8 31.96 74 85.8 38.39 73 30.5% -21.00 [-32.43, -9.57]
Total (95% CI) 282 276 100.0% -22.17 [-29.62, -14.72]
Heterogeneity: Tau² = 13.34; Chi² = 3.90, df = 3 (P = 0.27); I² = 23%
Test for overall effect: Z = 5.83 (P < 0.00001)
-50 -25 0 25 50
Favours 3 mg Doxepin Favours Placebo

Figure S57 – Meta-analysis of data for subjectively-determined quality of sleep in response to doxepin 3 mg

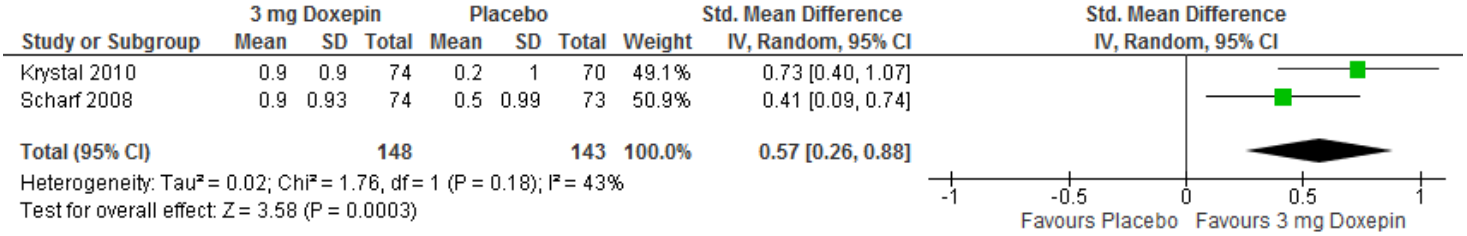

3 mg Doxepin
Placebo
Std. Mean Difference
Std. Mean Difference
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI
IV, Random, 95% CI
Krystal 2010 0.9 0.9 74 0.2 1 70 49.1% 0.73 [0.40, 1.07]
Scharf 2008 0.9 0.93 74 0.5 0.99 73 50.9% 0.41 [0.09, 0.74]
Total (95% CI) 148 143 100.0% 0.57 [0.26, 0.88]
Heterogeneity: Tau² = 0.02; Chi² = 1.76, df = 1 (P = 0.18); I² = 43%
Test for overall effect: Z = 3.58 (P = 0.0003)
-1 -0.5 0 0.5 1
Favours Placebo Favours 3 mg Doxepin

Figure S58 – Meta-analysis of data for PSG-determined sleep efficiency in response to doxepin 3 mg

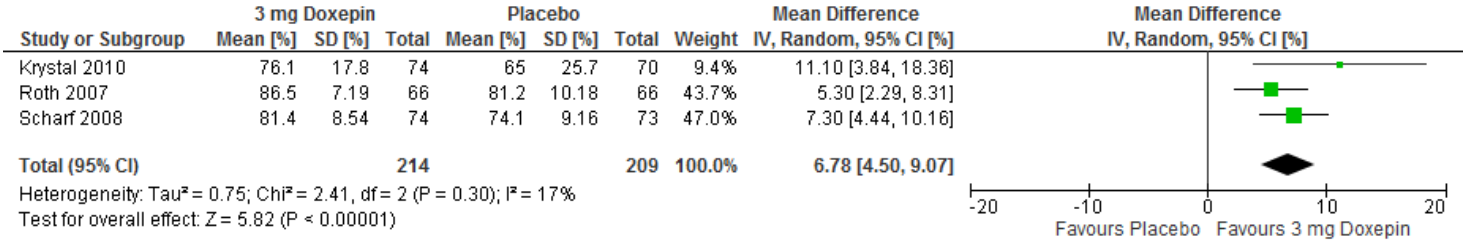

3 mg Doxepin
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [%] SD [%] Total Mean [%] SD [%] Total Weight IV, Random, 95% CI [%]
IV, Random, 95% CI [%]
Krystal 2010 76.1 17.8 74 65 25.7 70 9.4% 11.10 [3.84, 18.36]
Roth 2007 86.5 7.19 66 81.2 10.18 66 43.7% 5.30 [2.29, 8.31]
Scharf 2008 81.4 8.54 74 74.1 9.16 73 47.0% 7.30 [4.44, 10.16]
Total (95% CI) 214 209 100.0% 6.78 [4.50, 9.07]
Heterogeneity: Tau² = 0.75; Chi² = 2.41, df = 2 (P = 0.30); I² = 17%
Test for overall effect: Z = 5.82 (P < 0.00001)
-20 -10 0 10 20
Favours Placebo Favours 3 mg Doxepin

Figure S59 – Meta-analysis of data for PSG-determined number of awakenings in response to doxepin 3 mg

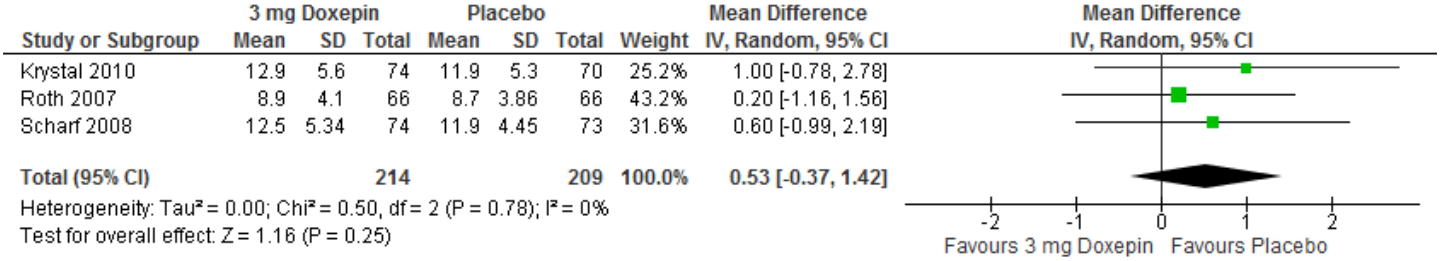

3 mg Doxepin
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI
IV, Random, 95% CI
Krystal 2010 12.9 5.6 74 11.9 5.3 70 25.2% 1.00 [-0.78, 2.78]
Roth 2007 8.9 4.1 66 8.7 3.86 66 43.2% 0.20 [-1.16, 1.56]
Scharf 2008 12.5 5.34 74 11.9 4.45 73 31.6% 0.60 [-0.99, 2.19]
Total (95% CI) 214 209 100.0% 0.53 [-0.37, 1.42]
Heterogeneity: Tau² = 0.00; Chi² = 0.50, df = 2 (P = 0.78); I² = 0%
Test for overall effect: Z = 1.16 (P = 0.25)
-2 -1 0 1 2
Favours 3 mg Doxepin Favours Placebo

Figure S60 – Meta-analysis of data for the occurrence of headache in response to doxepin 3 mg

Study or Subgroup	Doxepin 3 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Krystal 2010	4	70	10	70	19.2%	-0.09 [-0.18, 0.01]
Krystal 2011	3	68	7	67	24.0%	-0.06 [-0.15, 0.03]
Roth 2007	0	66	3	66	56.9%	-0.05 [-0.10, 0.01]
Total (95% CI)		**204**		**203**	**100.0%**	**-0.06 [-0.10, -0.01]**
Total events	7		20			

Heterogeneity: Tau² = 0.00; Chi² = 0.61, df = 2 (P = 0.74); I² = 0%
Test for overall effect: Z = 2.58 (P = 0.010)

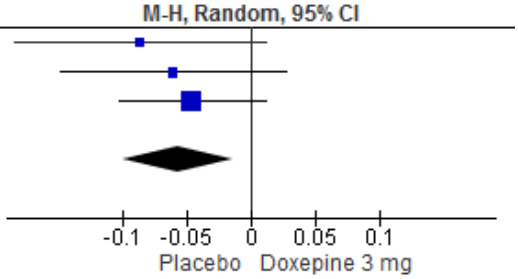


Figure S61 – Meta-analysis of data for the occurrence of somnolence in response to doxepin 3 mg

Study or Subgroup	Doxepin 3 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Krystal 2010	1	70	3	70	33.1%	-0.03 [-0.08, 0.03]
Krystal 2011	6	68	3	67	16.4%	0.04 [-0.04, 0.13]
Roth 2007	1	66	0	66	50.5%	0.02 [-0.03, 0.06]
Total (95% CI)		**204**		**203**	**100.0%**	**0.01 [-0.03, 0.04]**
Total events	8		6			

Heterogeneity: Tau² = 0.00; Chi² = 2.55, df = 2 (P = 0.28); I² = 21%
Test for overall effect: Z = 0.29 (P = 0.77)

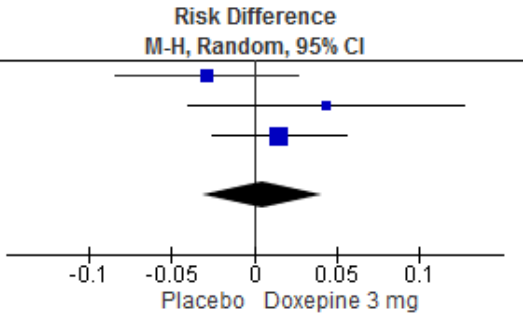


Figure S62 – Meta-analysis of data for the occurrence of diarrhea in response to doxepin 3 mg

Study or Subgroup	Doxepin 3 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Krystal 2010	0	70	1	70	51.5%	-0.01 [-0.05, 0.02]
Krystal 2011	1	68	0	67	48.5%	0.01 [-0.03, 0.05]
Total (95% CI)		**138**		**137**	**100.0%**	**-0.00 [-0.03, 0.03]**
Total events	1		1			

Heterogeneity: Tau² = 0.00; Chi² = 1.04, df = 1 (P = 0.31); I² = 4%
Test for overall effect: Z = 0.02 (P = 0.99)

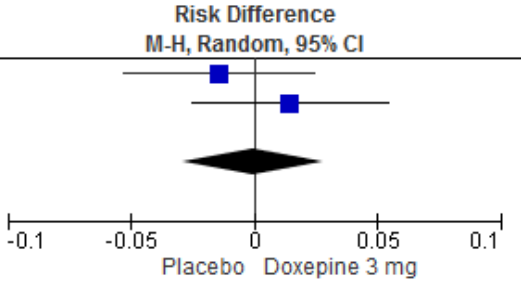


Figure S63 – Meta-analysis of data for the occurrence of upper respiratory tract infection in response to doxepin 3 mg

Study or Subgroup	Doxepin 3 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Krystal 2010	1	70	1	70	61.4%	0.00 [-0.04, 0.04]
Krystal 2011	2	68	1	67	38.6%	0.01 [-0.04, 0.06]
Total (95% CI)		**138**		**137**	**100.0%**	**0.01 [-0.03, 0.04]**
Total events	3		2			

Heterogeneity: Tau² = 0.00; Chi² = 0.21, df = 1 (P = 0.65); I² = 0%
Test for overall effect: Z = 0.36 (P = 0.72)

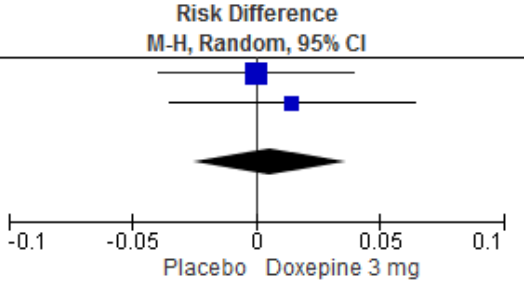


Table S15 – Summary of Findings table for doxepin 3 mg for the treatment of chronic insomnia

References: Krystal 2010(A); Krystal 2011(B); Roth 2007(C); Scharf 2008(D)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 3 mg Doxepin vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep latency in the doxepin groups was **2.3 minutes lower** (6.22 lower to 1.62 higher)	558 (4 studies)[A,B,C,D]
Sleep Latency (Subjective)	⊕⊕⊖⊖ **low**[1,6]	The mean sleep latency in the doxepin groups was **9.35 minutes lower** (21.89 lower to 3.19 higher)	291 (2 studies)[A,D]
Total Sleep Time* (PSG)	⊕⊕⊖⊖ **low**[1,3]	The mean total sleep time in the doxepin groups was **26.14 minutes higher** (18.49 to 33.79 higher)	558 (4 studies)[A,B,C,D]
Total Sleep Time (Subjective)	⊕⊖⊖⊖ **very low**[1,7,8]	The mean total sleep time in the doxepin groups was **43.57 minutes higher** (5.16 to 81.98 higher)	291 (2 studies)[A,D]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean wake after sleep onset in the doxepin groups was **22.17 minutes lower** (29.62 to 14.72 lower)	558 (4 studies)[A,B,C,D]
Wake After Sleep Onset (Subjective)	⊕⊕⊖⊖ **low**[1,9]	The mean wake after sleep onset in the doxepin group was **20.0 minutes lower** (39.07 to 0.92 lower)	147 (1 study)[D]
Quality of Sleep* (Subjective)	⊕⊕⊖⊖ **low**[1,5]	The mean quality of sleep in the doxepin groups was **0.57 standard deviations higher** (0.26 to 0.88 higher)	291 (2 studies)[A,D]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[1,4]	The mean sleep efficiency in the doxepin groups was **6.78 percent higher** (4.5 to 9.07 higher)	423 (3 studies)[A,C,D]
Number of Awakenings (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean number of awakenings in the doxepin groups was **0.53 awakenings higher** (0.37 lower to 1.42 higher)	423 (3 studies)[A,C,D]

* Critical Outcome, used to determine Quality of Evidence
[1] All studies funded by Industry
[2] 95% CI (-29.62, -14.72) crosses Clinical Significance (20 min)
[3] 95% CI (18.49, 33.79) crosses Clinical Significance (20 min)
[4] 95% CI (4.50, 9.07) crosses Clinical Significance (5%)
[5] 95% CI (0.26, 0.88) crosses Clinical Significance (SMD 0.5)
[6] 95% CI (-21.89, 3.19) crosses Clinical Significance (20 min)
[7] Heterogeneity (I^2 = 82%) greater than allowance (75%)
[8] 95% CI (5.16, 81.98) crosses Clinical Significance (30 min)
[9] 95% CI (-39.07, -0.92) crosses Clinical Significance (30 min)

Figure S64 – Meta-analysis of data for PSG-determined sleep latency in response to doxepin 6 mg

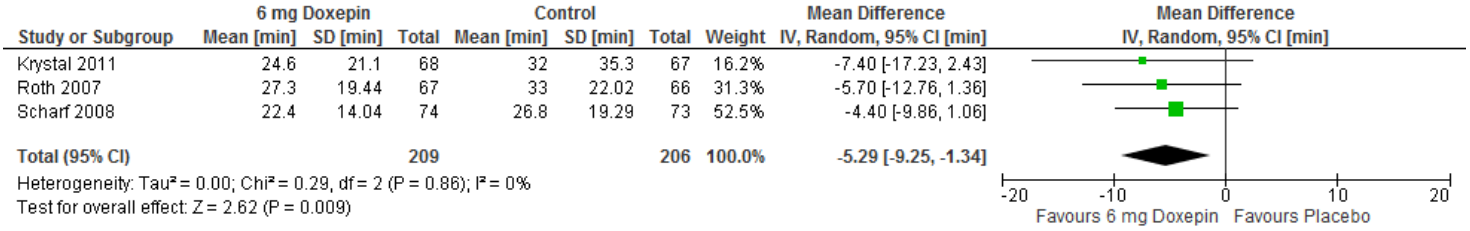

6 mg Doxepin
Control
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Krystal 2011 24.6 21.1 68 32 35.3 67 16.2% -7.40 [-17.23, 2.43]
Roth 2007 27.3 19.44 67 33 22.02 66 31.3% -5.70 [-12.76, 1.36]
Scharf 2008 22.4 14.04 74 26.8 19.29 73 52.5% -4.40 [-9.86, 1.06]
Total (95% CI) 209 206 100.0% -5.29 [-9.25, -1.34]
Heterogeneity: Tau² = 0.00; Chi² = 0.29, df = 2 (P = 0.86); I² = 0%
Test for overall effect: Z = 2.62 (P = 0.009)
-20 -10 0 10 20
Favours 6 mg Doxepin Favours Placebo

Figure S65 – Meta-analysis of data for PSG-determined total sleep time in response to doxepin 6 mg

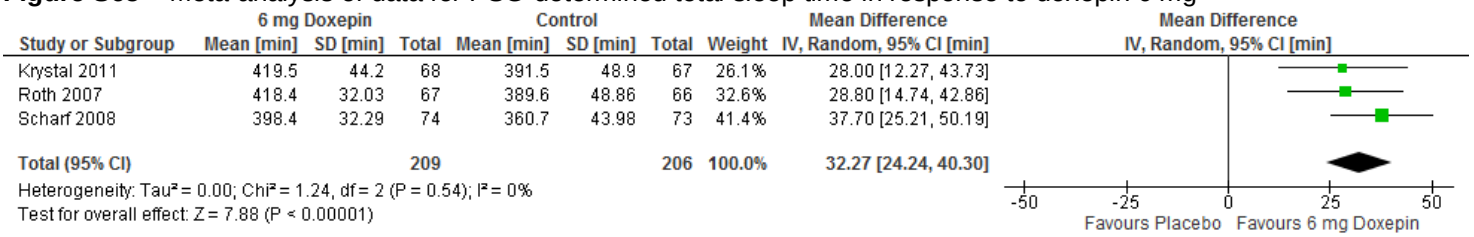

6 mg Doxepin
Control
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Krystal 2011 419.5 44.2 68 391.5 48.9 67 26.1% 28.00 [12.27, 43.73]
Roth 2007 418.4 32.03 67 389.6 48.86 66 32.6% 28.80 [14.74, 42.86]
Scharf 2008 398.4 32.29 74 360.7 43.98 73 41.4% 37.70 [25.21, 50.19]
Total (95% CI) 209 206 100.0% 32.27 [24.24, 40.30]
Heterogeneity: Tau² = 0.00; Chi² = 1.24, df = 2 (P = 0.54); I² = 0%
Test for overall effect: Z = 7.88 (P < 0.00001)
-50 -25 0 25 50
Favours Placebo Favours 6 mg Doxepin

Figure S66 – Meta-analysis of data for subjectively-determined total sleep time in response to doxepin 6 mg

Study or Subgroup	6 mg Doxepin Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Lankford 2012	346.1	66.4	130	336.4	64.7	124	56.7%	9.70 [-6.42, 25.82]
Scharf 2008	370.8	64.59	74	340	71.79	73	43.3%	30.80 [8.71, 52.89]
Total (95% CI)			**204**			**197**	**100.0%**	**18.84 [-1.65, 39.34]**

Heterogeneity: Tau² = 125.28; Chi² = 2.29, df = 1 (P = 0.13); I² = 56%
Test for overall effect: Z = 1.80 (P = 0.07)

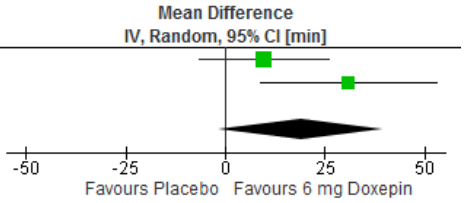


Figure S67 – Meta-analysis of data for PSG-determined wake after sleep onset in response to doxepin 6 mg

Study or Subgroup	6 mg Doxepin Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Krystal 2011	40.7	37.3	68	60.5	38.8	67	29.2%	-19.80 [-32.64, -6.96]
Roth 2007	38.1	25.16	67	61.1	45.79	66	30.4%	-23.00 [-35.58, -10.42]
Scharf 2008	59.5	28.3	74	85.8	38.39	73	40.4%	-26.30 [-37.21, -15.39]
Total (95% CI)			**209**			**206**	**100.0%**	**-23.40 [-30.34, -16.46]**

Heterogeneity: Tau² = 0.00; Chi² = 0.58, df = 2 (P = 0.75); I² = 0%
Test for overall effect: Z = 6.61 (P < 0.00001)

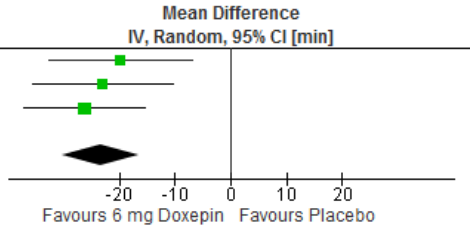


Figure S68 – Meta-analysis of data for subjectively-determined wake after sleep onset in response to doxepin 6 mg

Study or Subgroup	6 mg Doxepin Mean [min]	SD [min]	Total	Control Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Lankford 2012	66.5	43.9	130	78.9	56.5	124	70.3%	-12.40 [-24.88, 0.08]
Scharf 2008	70.2	57.08	74	89.3	61.56	73	29.7%	-19.10 [-38.30, 0.10]
Total (95% CI)			**204**			**197**	**100.0%**	**-14.39 [-24.86, -3.93]**

Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.57); I² = 0%
Test for overall effect: Z = 2.70 (P = 0.007)

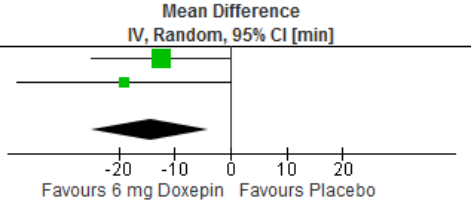


Figure S69 – Meta-analysis of data for subjectively-determined quality of sleep in response to doxepin 6 mg

Study or Subgroup	6 mg Doxepin Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Lankford 2012	0.4	1	130	0.2	1.1	127	61.9%	0.19 [-0.06, 0.43]
Scharf 2008	0.9	0.92	74	0.5	0.99	73	38.1%	0.42 [0.09, 0.74]
Total (95% CI)			**204**			**200**	**100.0%**	**0.28 [0.06, 0.49]**

Heterogeneity: Tau² = 0.00; Chi² = 1.18, df = 1 (P = 0.28); I² = 15%
Test for overall effect: Z = 2.51 (P = 0.01)

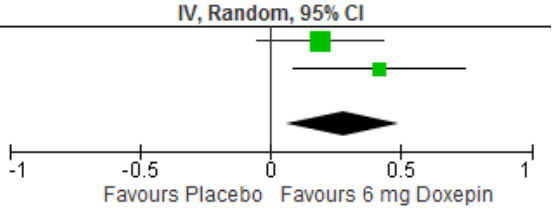


Figure S70 – Meta-analysis of data for PSG-determined sleep efficiency in response to doxepin 6 mg

Study or Subgroup	6 mg Doxepin Mean [%]	SD [%]	Total	Control Mean [%]	SD [%]	Total	Weight	Mean Difference IV, Random, 95% CI [%]
Roth 2007	87.2	6.67	67	81.2	10.18	66	44.1%	6.00 [3.07, 8.93]
Scharf 2008	83	6.73	74	75.1	9.16	73	55.9%	7.90 [5.30, 10.50]
Total (95% CI)			**141**			**139**	**100.0%**	**7.06 [5.12, 9.01]**

Heterogeneity: Tau² = 0.00; Chi² = 0.90, df = 1 (P = 0.34); I² = 0%
Test for overall effect: Z = 7.12 (P < 0.00001)

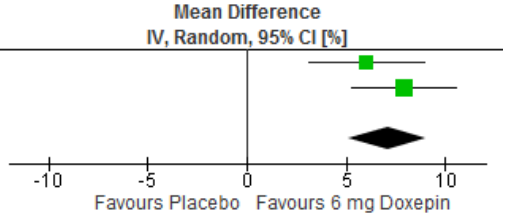


Figure S71 – Meta-analysis of data for PSG-determined number of awakenings in response to doxepin 6 mg

Study or Subgroup	6 mg Doxepin Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Roth 2007	9	4.1	67	8.7	3.86	66	54.8%	0.30 [-1.05, 1.65]
Scharf 2008	12.5	4.76	74	11.9	4.45	73	45.2%	0.60 [-0.89, 2.09]
Total (95% CI)			**141**			**139**	**100.0%**	**0.44 [-0.57, 1.44]**

Heterogeneity: Tau² = 0.00; Chi² = 0.09, df = 1 (P = 0.77); I² = 0%
Test for overall effect: Z = 0.85 (P = 0.39)

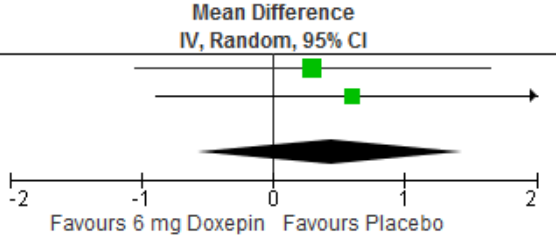


Figure S72 – Meta-analysis of data for the occurrence of headache in response to doxepin 6 mg

Study or Subgroup	Doxepin 6 mg Events	Doxepin 6 mg Total	Placebo Events	Placebo Total	Weight	Risk Difference M-H, Random, 95% CI
Krystal 2011	0	68	6	67	15.8%	-0.09 [-0.16, -0.02]
Lankford 2012	0	130	5	124	59.5%	-0.04 [-0.08, -0.00]
Roth 2007	1	66	3	66	24.6%	-0.03 [-0.09, 0.03]
Total (95% CI)		**264**		**257**	**100.0%**	**-0.05 [-0.07, -0.02]**
Total events	1		14			

Heterogeneity: Tau² = 0.00; Chi² = 1.86, df = 2 (P = 0.40); I² = 0%
Test for overall effect: Z = 3.09 (P = 0.002)

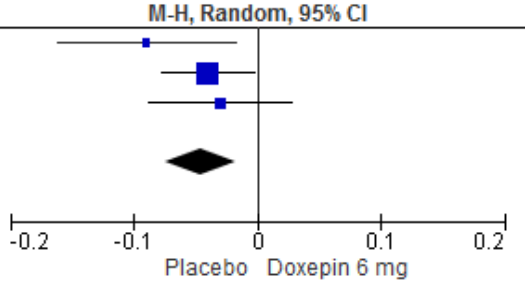


Figure S73 – Meta-analysis of data for the occurrence of somnolence in response to doxepin 6 mg

Study or Subgroup	Doxepin 6 mg Events	Doxepin 6 mg Total	Placebo Events	Placebo Total	Weight	Risk Difference M-H, Random, 95% CI
Krystal 2011	5	68	3	67	18.6%	0.03 [-0.05, 0.11]
Lankford 2012	12	130	4	124	34.1%	0.06 [0.00, 0.12]
Roth 2007	2	66	0	66	47.3%	0.03 [-0.02, 0.08]
Total (95% CI)		**264**		**257**	**100.0%**	**0.04 [0.01, 0.07]**
Total events	19		7			

Heterogeneity: Tau² = 0.00; Chi² = 0.73, df = 2 (P = 0.69); I² = 0%
Test for overall effect: Z = 2.30 (P = 0.02)

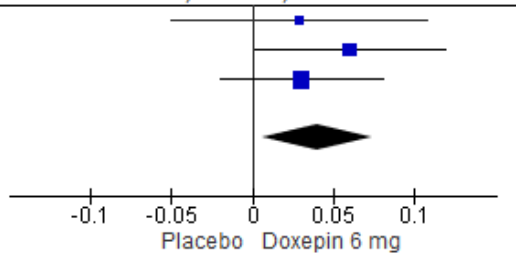


Table S16 – Summary of Findings table for doxepin 6 mg for the treatment of chronic insomnia

References: Krystal 2011(A); Roth 2007(B); Lankford 2012(C); Scharf 2008(D)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 6 mg Doxepin vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep latency in the doxepin groups was **5.29 minutes lower** (9.25 to 1.34 lower)	415 (3 studies)[A,B,D]
Total Sleep Time* (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean total sleep time in the doxepin groups was **32.27 minutes higher** (24.24 to 40.3 higher)	415 (3 studies)[A,B,D]
Total Sleep Time (Subjective)	⊕⊕⊖⊖ **low**[1,3]	The mean total sleep time in the doxepin groups was **18.84 minutes higher** (1.65 lower to 39.34 higher)	401 (2 studies)[C,D]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean wake after sleep onset in the doxepin groups was **23.4 minutes lower** (30.34 to 16.46 lower)	415 (3 studies)[A,B,D]
Wake After Sleep Onset (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean wake after sleep onset in the doxepin groups was **14.39 minutes lower** (24.86 to 3.93 lower)	401 (2 studies)[C,D]
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean quality of sleep in the doxepin groups was **0.28 standard deviations higher** (0.06 to 0.49 higher)	404 (2 studies)[C,D]
Sleep Efficiency (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep efficiency in the doxepin groups was **7.06 percent higher** (5.12 to 9.01 higher)	280 (2 studies)[B,D]
Number of Awakenings (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean number of awakenings in the doxepin groups was **0.44 awakenings higher** (0.57 lower to 1.44 higher)	280 (2 studies)[B,D]

* Critical Outcome, used to determine Quality of Evidence
[1] All studies funded by industry
[2] 95% CI (-30.34, -16.46) crosses Clinical Significance (20 min)
[3] 95% CI (-1.65, 39.34) crosses Clinical Significance (30 min)

Trazadone - Summary of Findings Table

Table S17 – Summary of Findings table for trazodone 50 mg for the treatment of chronic insomnia

References: Walsh 1998(A)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 50 mg Trazadone vs Placebo	No of Participants (studies)
Sleep Latency* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep latency in the trazadone group was **10.20 minutes lower** (11.44 to 8.95 lower)	187 (1 study)[A]
Total Sleep Time* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean total sleep time in the trazadone group was **21.80 minutes higher** (20.10 to 23.49 higher)	187 (1 study)[A]
Wake After Sleep Onset* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean wake after sleep onset in the trazadone group was **7.70 minutes lower** (8.89 to 6.5 lower)	187 (1 study)[A]
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean quality of sleep in the trazadone group was **0.13 points**[2] **lower** (0.14 to 0.11 lower)	187 (1 study)[A]
Number of Awakenings (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean number of awakenings in the trazadone group was **0.40 awakenings lower** (0.42 to 0.37 lower)	187 (1 study)[A]

* Critical Outcome, used to determine Quality of Evidence
[1] Study funded by industry
[2] 4-point scale (1=Excellent, 4=Poor)

Tiagabine - Meta-Analyses and Summary of Findings Tables

Figure S74 – Meta-analysis of data for PSG-determined sleep latency in response to tiagabine 4 mg

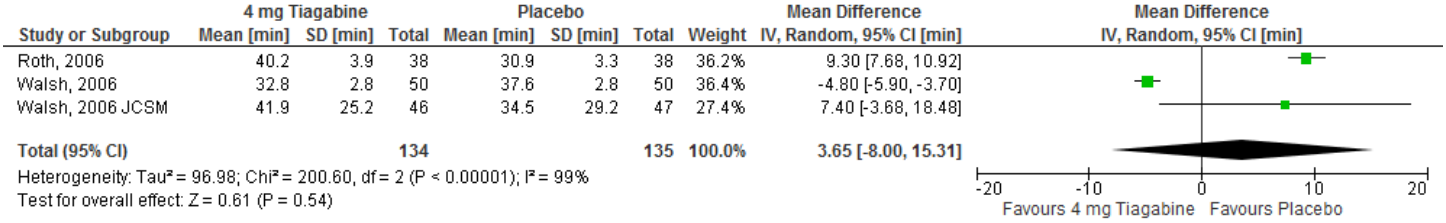

4 mg Tiagabine
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Roth, 2006 40.2 3.9 38 30.9 3.3 38 36.2% 9.30 [7.68, 10.92]
Walsh, 2006 32.8 2.8 50 37.6 2.8 50 36.4% -4.80 [-5.90, -3.70]
Walsh, 2006 JCSM 41.9 25.2 46 34.5 29.2 47 27.4% 7.40 [-3.68, 18.48]
Total (95% CI) 134 135 100.0% 3.65 [-8.00, 15.31]
Heterogeneity: Tau² = 96.98; Chi² = 200.60, df = 2 (P < 0.00001); I² = 99%
Test for overall effect: Z = 0.61 (P = 0.54)
-20 -10 0 10 20
Favours 4 mg Tiagabine Favours Placebo

Figure S75 – Meta-analysis of data for subjectively-determined sleep latency in response to tiagabine 4 mg

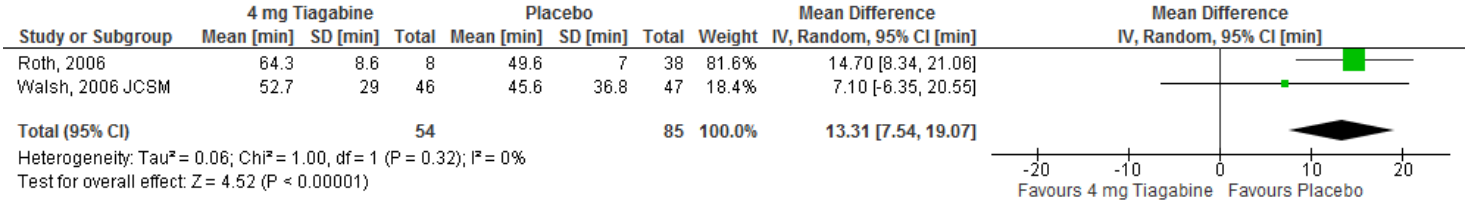

4 mg Tiagabine
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Roth, 2006 64.3 8.6 8 49.6 7 38 81.6% 14.70 [8.34, 21.06]
Walsh, 2006 JCSM 52.7 29 46 45.6 36.8 47 18.4% 7.10 [-6.35, 20.55]
Total (95% CI) 54 85 100.0% 13.31 [7.54, 19.07]
Heterogeneity: Tau² = 0.06; Chi² = 1.00, df = 1 (P = 0.32); I² = 0%
Test for overall effect: Z = 4.52 (P < 0.00001)
-20 -10 0 10 20
Favours 4 mg Tiagabine Favours Placebo

Figure S76 – Meta-analysis of data for PSG-determined total sleep time in response to tiagabine 4 mg

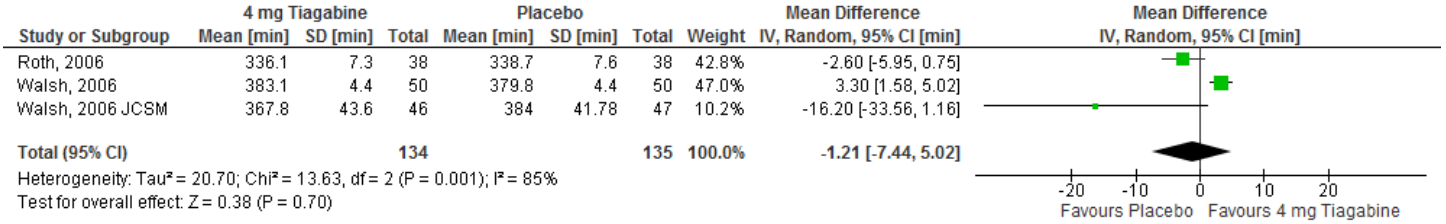

4 mg Tiagabine
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Roth, 2006 336.1 7.3 38 338.7 7.6 38 42.8% -2.60 [-5.95, 0.75]
Walsh, 2006 383.1 4.4 50 379.8 4.4 50 47.0% 3.30 [1.58, 5.02]
Walsh, 2006 JCSM 367.8 43.6 46 384 41.78 47 10.2% -16.20 [-33.56, 1.16]
Total (95% CI) 134 135 100.0% -1.21 [-7.44, 5.02]
Heterogeneity: Tau² = 20.70; Chi² = 13.63, df = 2 (P = 0.001); I² = 85%
Test for overall effect: Z = 0.38 (P = 0.70)
-20 -10 0 10 20
Favours Placebo Favours 4 mg Tiagabine

Figure S77 – Meta-analysis of data for subjectively-determined total sleep time in response to tiagabine 4 mg

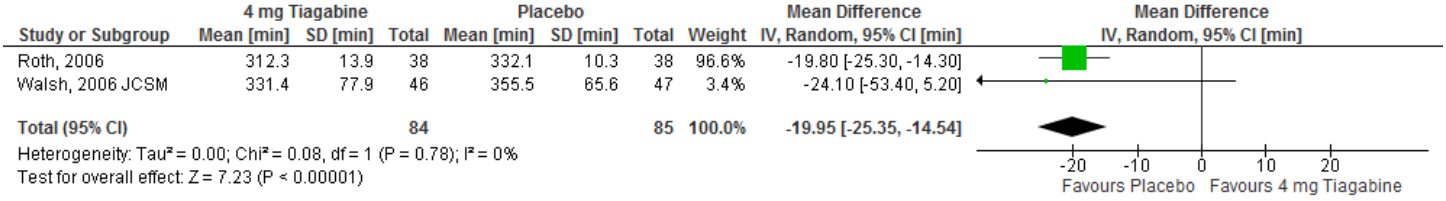

4 mg Tiagabine
Placebo
Mean Difference
Mean Difference
Study or Subgroup Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Roth, 2006 312.3 13.9 38 332.1 10.3 38 96.6% -19.80 [-25.30, -14.30]
Walsh, 2006 JCSM 331.4 77.9 46 355.5 65.6 47 3.4% -24.10 [-53.40, 5.20]
Total (95% CI) 84 85 100.0% -19.95 [-25.35, -14.54]
Heterogeneity: Tau² = 0.00; Chi² = 0.08, df = 1 (P = 0.78); I² = 0%
Test for overall effect: Z = 7.23 (P < 0.00001)
-20 -10 0 10 20
Favours Placebo Favours 4 mg Tiagabine

Figure S78 – Meta-analysis of data for PSG-determined wake after sleep onset in response to tiagabine 4 mg

Study or Subgroup	4 mg Tiagabine Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Roth, 2006	112.5	6.6	38	118.2	6.6	38	40.7%	-5.70 [-8.67, -2.73]
Walsh, 2006	68.2	3.9	50	66.9	3.9	50	43.6%	1.30 [-0.23, 2.83]
Walsh, 2006 JCSM	76.1	34	46	68.5	27.8	47	15.7%	7.60 [-5.04, 20.24]
Total (95% CI)			**134**			**135**	**100.0%**	**-0.56 [-6.77, 5.65]**

Heterogeneity: Tau² = 22.39; Chi² = 18.33, df = 2 (P = 0.0001); I² = 89%
Test for overall effect: Z = 0.18 (P = 0.86)

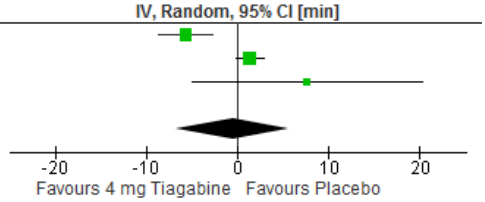


Figure S79 – Meta-analysis of data for subjectively-determined wake after sleep onset in response to tiagabine 4 mg

Study or Subgroup	4 mg Tiagabine Mean [min]	SD [min]	Total	Placebo Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Roth, 2006	97.6	9.5	38	93.2	10.8	38	97.1%	4.40 [-0.17, 8.97]
Walsh, 2006 JCSM	75.4	62.8	46	74.9	66.3	47	2.9%	0.50 [-25.74, 26.74]
Total (95% CI)			**84**			**85**	**100.0%**	**4.29 [-0.22, 8.79]**

Heterogeneity: Tau² = 0.00; Chi² = 0.08, df = 1 (P = 0.77); I² = 0%
Test for overall effect: Z = 1.86 (P = 0.06)

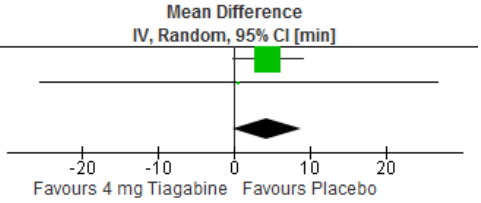


Figure S80 – Meta-analysis of data for subjectively-determined quality of sleep in response to tiagabine 4 mg

Study or Subgroup	4 mg Tiagabine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Roth, 2006	2.7	0.1	38	2.6	0.1	38	49.2%	0.99 [0.51, 1.47]
Walsh, 2006 JCSM	56.9	21.8	46	57.1	20.7	47	50.8%	-0.01 [-0.42, 0.40]
Total (95% CI)			**84**			**85**	**100.0%**	**0.48 [-0.50, 1.46]**

Heterogeneity: Tau² = 0.45; Chi² = 9.74, df = 1 (P = 0.002); I² = 90%
Test for overall effect: Z = 0.97 (P = 0.33)

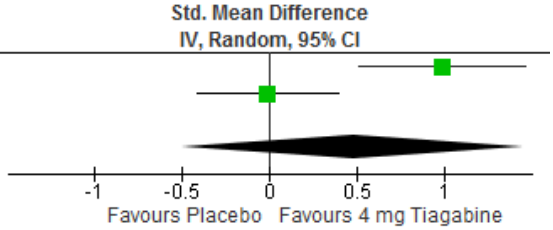


Figure S81 – Meta-analysis of data for PSG-determined sleep efficiency in response to tiagabine 4 mg

Study or Subgroup	4 mg Tiagabine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Roth, 2006	70	1.5	38	70.6	1.6	38	36.3%	-0.60 [-1.30, 0.10]
Walsh, 2006	85.6	1.1	50	86	1.1	50	61.8%	-0.40 [-0.83, 0.03]
Walsh, 2006 JCSM	76.6	9.1	46	80	8.7	47	2.0%	-3.40 [-7.02, 0.22]
Total (95% CI)			**134**			**135**	**100.0%**	**-0.53 [-1.05, -0.02]**

Heterogeneity: Tau² = 0.06; Chi² = 2.75, df = 2 (P = 0.25); I² = 27%
Test for overall effect: Z = 2.03 (P = 0.04)

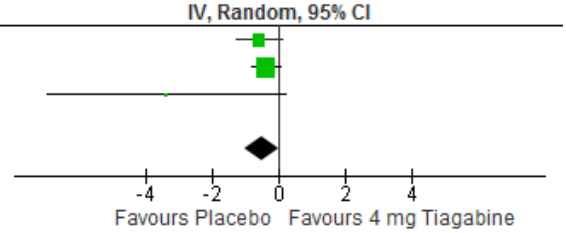


Figure S82 – Meta-analysis of data for PSG-determined number of awakenings in response to tiagabine 4 mg

Study or Subgroup	4 mg Tiagabine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Walsh, 2006	10.6	5	50	9.8	5	50	83.5%	0.80 [-1.16, 2.76]
Walsh, 2006 JCSM	31.1	12	46	32.1	9.5	47	16.5%	-1.00 [-5.40, 3.40]
Total (95% CI)			**96**			**97**	**100.0%**	**0.50 [-1.29, 2.29]**

Heterogeneity: Tau² = 0.00; Chi² = 0.54, df = 1 (P = 0.46); I² = 0%
Test for overall effect: Z = 0.55 (P = 0.58)

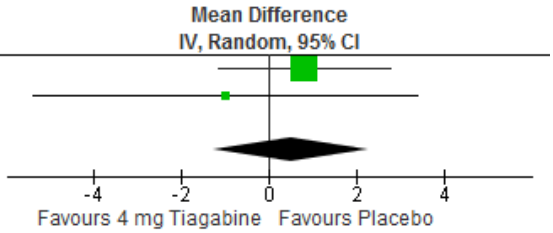


Figure S83 – Meta-analysis of data for subjectively-determined number of awakenings in response to tiagabine 4 mg

Study or Subgroup	4 mg Tiagabine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Roth, 2006	3.4	0.3	38	3.4	0.4	38	73.7%	0.00 [-0.16, 0.16]
Walsh, 2006 JCSM	3.4	1.9	46	4.2	3.3	47	26.3%	-0.80 [-1.89, 0.29]
Total (95% CI)			**84**			**85**	**100.0%**	**-0.21 [-0.90, 0.48]**

Heterogeneity: Tau² = 0.16; Chi² = 2.02, df = 1 (P = 0.16); I² = 51%
Test for overall effect: Z = 0.60 (P = 0.55)

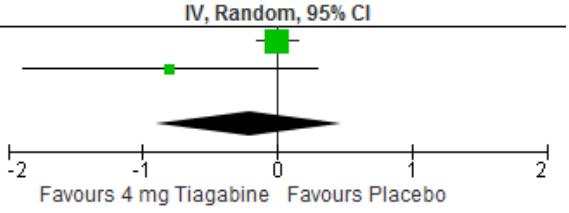


Figure S84 – Meta-analysis of data for the occurrence of headache in response to tiagabine 4 mg

Study or Subgroup	Tiagabine 4 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Roth 2006	1	38	1	38	43.6%	0.00 [-0.07, 0.07]
Walsh JCSM 2006	4	46	2	47	22.7%	0.04 [-0.06, 0.14]
Walsh SLEEP 2006	2	51	3	53	33.7%	-0.02 [-0.10, 0.06]
Total (95% CI)		**135**		**138**	**100.0%**	**0.00 [-0.04, 0.05]**
Total events	7		6			

Heterogeneity: Tau² = 0.00; Chi² = 0.93, df = 2 (P = 0.63); I² = 0%
Test for overall effect: Z = 0.17 (P = 0.86)

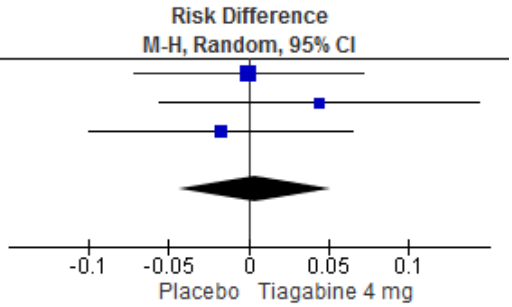


Figure S85 – Meta-analysis of data for the occurrence of nausea in response to tiagabine 4 mg

Study or Subgroup	Tiagabine 4 mg Events	Total	Placebo Events	Total	Weight	Risk Difference M-H, Random, 95% CI
Roth 2006	1	38	0	38	16.5%	0.03 [-0.04, 0.10]
Walsh JCSM 2006	1	46	0	47	24.1%	0.02 [-0.04, 0.08]
Walsh SLEEP 2006	0	51	0	53	59.4%	0.00 [-0.04, 0.04]
Total (95% CI)		**135**		**138**	**100.0%**	**0.01 [-0.02, 0.04]**
Total events	2		0			

Heterogeneity: Tau² = 0.00; Chi² = 0.78, df = 2 (P = 0.68); I² = 0%
Test for overall effect: Z = 0.66 (P = 0.51)

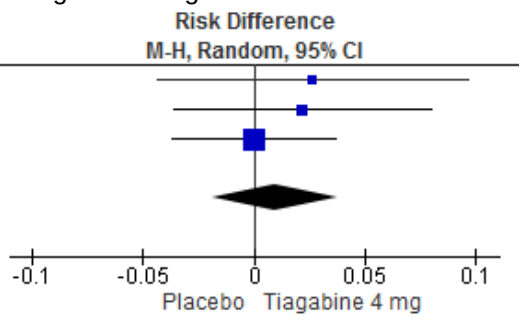


Table S18 – Summary of Findings table for tiagabine 4 mg for the treatment of chronic insomnia

References: Roth 2006(A); Walsh 2006(B); Walsh 2006 JCSM(C)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 4 mg Tiagabine vs Placebo	No of Participants (studies)
Sleep Latency (PSG)	⊕⊖⊖⊖ **very low**[1,2,3]	The mean sleep latency in the tiagabine groups was **3.65 minutes higher** (8 lower to 15.31 higher)	269 (3 studies)[A,B,C]
Sleep Latency (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean sleep latency in the tiagabine groups was **13.31 minutes higher** (7.54 to 19.07 higher)	139 (2 studies)[A,C]
Total Sleep Time* (PSG)	⊕⊕⊖⊖ **low**[3,5]	The mean total sleep time in the tiagabine groups was **1.21 minutes lower** (7.44 lower to 5.02 higher)	269 (3 studies)[A,B,C]
Total Sleep Time (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean total sleep time in the tiagabine groups was **19.95 minutes lower** (25.35 to 14.54 lower)	169 (2 studies)[A,C]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[3,4]	The mean wake after sleep onset in the tiagabine groups was **0.56 minutes lower** (6.77 lower to 5.65 higher)	269 (3 studies)[A,B,C]
Wake After Sleep Onset (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean wake after sleep onset in the tiagabine groups was **4.29 minutes higher** (0.22 lower to 8.79 higher)	169 (2 studies)[A,C]
Quality of Sleep* (Subjective)	⊕⊖⊖⊖ **very low**[3,7,8]	The mean quality of sleep in the tiagabine groups was **0.48 standard deviations higher** (0.5 lower to 1.46 higher)	169 (2 studies)[A,C]
Sleep Efficiency (PSG)	⊕⊕⊕⊖ **moderate**[3]	The mean sleep efficiency in the tiagabine groups was **0.53 percent lower** (1.05 to 0.02 lower)	269 (3 studies)[A,B,C]
Number of Awakenings (PSG)	⊕⊕⊖⊖ **low**[3,6]	The mean number of awakenings in the tiagabine groups was **0.5 awakenings higher** (1.29 lower to 2.29 higher)	193 (2 studies)[B,C]
Number of Awakenings (Subjective)	⊕⊕⊖⊖ **low**[3,9]	The mean number of awakenings in the tiagabine groups was **0.21 awakenings lower** (0.9 lower to 0.48 higher)	169 (2 studies)[A,C]

* Critical Outcome, used to determine Quality of Evidence
[1] Heterogeneity (I^2 = 99%) greater than allowance (75%)
[2] 95% CI (-8.0, 15.31) crosses Clinical Significance (10 min)
[3] All studies funded by industry
[4] Heterogeneity (I^2 = 89%) greater than allowance (75%)
[5] Heterogeneity (I^2 = 85%) greater than allowance (75%)
[6] 95% CI (-1.29, 2.29) crosses Clinical Significance (2 awakenings)
[7] Heterogeneity (I^2 = 90%) greater than allowance (75%)
[8] 95% CI (-0.50, 1.46) crosses zero standard mean difference
[9] 95% CI (-0.90, 0.48) crosses Clinical Significance (0.5 awakenings)

Table S19 – Summary of Findings table for tiagabine 6 mg for the treatment of chronic insomnia

References: Roth 2006(A); Walsh 2006 JCSM(B)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 6 mg Tiagabine vs Placebo	No of Participants (studies)
Sleep Latency (PSG)	⊕⊕⊖⊖ **low**[3,7]	The mean sleep latency in the tiagabine groups was **6.9 minutes higher** (2.22 to 11.58 higher)	175 (2 studies) [A,B]
Sleep Latency (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean sleep latency in the tiagabine groups was **5.68 minutes higher** (3.05 to 8.3 higher)	175 (2 studies) [A,B]
Total Sleep Time* (PSG)	⊕⊕⊕⊖ **moderate**[3]	The mean total sleep time in the tiagabine groups was **7.17 minutes higher** (0.26 lower to 14.59 higher)	175 (2 studies) [A,B]
Total Sleep Time (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean total sleep time in the tiagabine groups was **9.65 minutes lower** (14.05 to 5.25 lower)	175 (2 studies) [A,B]
Wake After Sleep Onset* (PSG)	⊕⊖⊖⊖ **very low**[1,2,3]	The mean wake after sleep onset in the tiagabine groups was **9.24 minutes lower** (24.78 lower to 6.3 higher)	175 (2 studies) [A,B]
Wake After Sleep Onset (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean wake after sleep onset in the tiagabine groups was **5.68 minutes higher** (3.05 to 8.3 higher)	175 (2 studies) [A,B]
Quality of Sleep* (Subjective)	⊕⊕⊖⊖ **low**[3,4]	The mean quality of sleep in the tiagabine groups was **0.01 standard deviations higher** (0.28 lower to 0.31 higher)	175 (2 studies) [A,B]
Sleep Efficiency (PSG)	⊕⊕⊕⊖ **moderate**[3]	The mean sleep efficiency in the tiagabine groups was **1.46 percent higher** (0.15 lower to 3.06 higher)	175 (2 studies) [A,B]
Number of Awakenings (Subjective)	⊕⊖⊖⊖ **very low**[3,5,6]	The mean number of awakenings in the tiagabine groups was **0.49 awakenings lower** (1.84 lower to 0.87 higher)	175 (2 studies) [A,B]

* Critical Outcome, used to determine Quality of Evidence
[1] Heterogeneity (I^2 = 81%) crosses threshold (75%)
[2] 95% CI (-24.78, 6.30) crosses Clinical Significance (20 min)
[3] All studies funded by industry
[4] 95% CI (-0.28, 0.31) crosses zero standard mean difference
[5] Heterogeneity (I^2 = 83%) crosses threshold (75%)
[6] 95% CI (-1.84, 0.87) crosses Clinical Significance (0.5 awakenings)
[7] 95% CI (2.22, 11.58) crosses Clinical Significance (10 min)

Table S20 – Summary of Findings table for tiagabine 8 mg for the treatment of chronic insomnia

References: Roth 2006(A); Walsh 2006(B); Walsh 2006 JCSM(C)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 6 mg Tiagabine vs Placebo	No of Participants (studies)
Sleep Latency (PSG)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep latency in the tiagabine groups was **1.22 minutes lower** (2.66 lower to 0.22 higher)	271 (3 studies) [A,B,C]
Sleep Latency (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean sleep latency in the tiagabine groups was **2.12 minutes lower** (3.48 to 0.76 lower)	171 (2 studies) [A,C]
Total Sleep Time* (PSG)	⊕⊕⊖⊖ **low**[1,3]	The mean total sleep time in the tiagabine groups was **3.49 minutes higher** (6.43 lower to 13.42 higher)	271 (3 studies) [A,B,C]
Total Sleep Time (Subjective)	⊕⊖⊖⊖ **very low**[1,7,8]	The mean total sleep time in the tiagabine groups was **16.09 minutes lower** (44.97 lower to 12.79 higher)	171 (2 studies) [A,C]
Wake After Sleep Onset* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean wake after sleep onset in the tiagabine groups was **2.42 minutes lower** (10.35 lower to 5.51 higher)	271 (3 studies) [A,B,C]
Wake After Sleep Onset (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean wake after sleep onset in the tiagabine groups was **9.71 minutes higher** (5.7 to 13.72 higher)	171 (2 studies) [A,C]
Quality of Sleep* (Subjective)	⊕⊖⊖⊖ **very low**[1,5,6]	The mean quality of sleep in the tiagabine groups was **0.37 standard deviations higher** (0.65 lower to 1.39 higher)	171 (2 studies) [A,C]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[1,3]	The mean sleep efficiency in the tiagabine groups was **0.68 percent higher** (1.41 lower to 2.76 higher)	271 (3 studies) [A,B,C]
Number of Awakenings (PSG)	⊕⊕⊖⊖ **low**[1,4]	The mean number of awakenings in the tiagabine groups was **0.88 awakenings lower** (3.7 lower to 1.95 higher)	192 (2 studies) [B,C]
Number of Awakenings (Subjective)	⊕⊕⊖⊖ **low**[1,9]	The mean number of awakenings in the tiagabine groups was **0.3 awakenings higher** (0.38 lower to 0.98 higher)	171 (2 studies) [A,C]

* Critical Outcome, used to determine Quality of Evidence
[1] All studies funded by industry
[2] Heterogeneity (I^2 = 93%) greater than allowance (75%)
[3] Heterogeneity (I^2 = 94%) greater than allowance (75%)
[4] 95% CI (-3.70, 1.95) crosses Clinical Significance (2 awakenings)
[5] Heterogeneity (I^2 = 91%) greater than allowance (75%)
[6] 95% CI (-0.65, 1.39) crosses zero standard mean difference
[7] Heterogeneity (I^2 = 89%) greater than allowance (75%)
[8] 95% CI (-44.97, 12.79) crosses Clinical Significance
[9] 95% CI (-0.38, 0.98) crosses Clinical Significance (0.5 awakenings)

Diphenhydramine - Meta-Analyses and Summary of Findings Table

Figure S86 – Meta-analysis of data for subjectively-determined sleep latency in response to diphenhydramine 50 mg

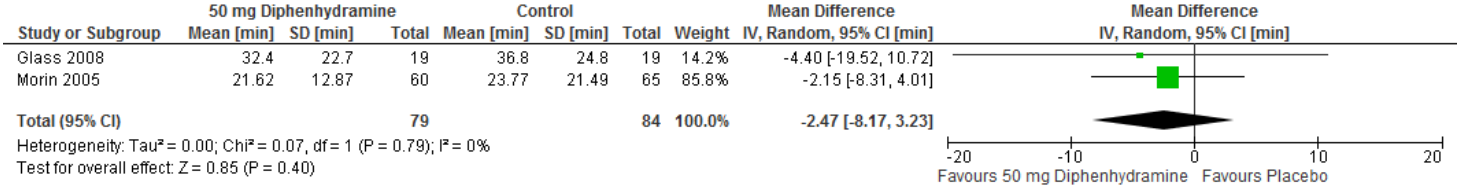

50 mg Diphenhydramine
Control
Mean Difference
Mean Difference
Study or Subgroup
Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Glass 2008 32.4 22.7 19 36.8 24.8 19 14.2% -4.40 [-19.52, 10.72]
Morin 2005 21.62 12.87 60 23.77 21.49 65 85.8% -2.15 [-8.31, 4.01]
Total (95% CI) 79 84 100.0% -2.47 [-8.17, 3.23]
Heterogeneity: Tau² = 0.00; Chi² = 0.07, df = 1 (P = 0.79); I² = 0%
Test for overall effect: Z = 0.85 (P = 0.40)
-20
-10
0
10
20
Favours 50 mg Diphenhydramine
Favours Placebo

Figure S87 – Meta-analysis of data for subjectively-determined total sleep time in response to diphenhydramine 50 mg

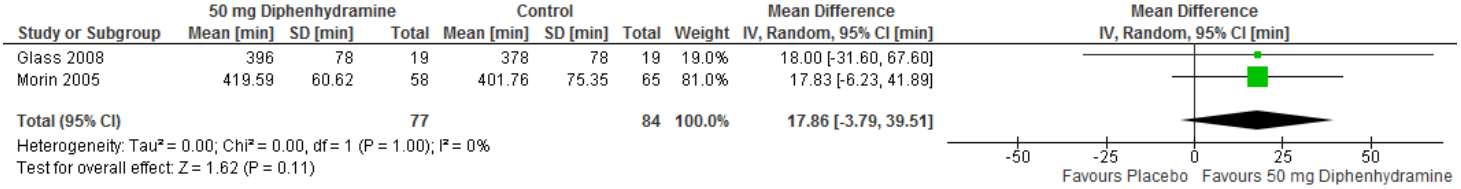

50 mg Diphenhydramine
Control
Mean Difference
Mean Difference
Study or Subgroup
Mean [min] SD [min] Total Mean [min] SD [min] Total Weight IV, Random, 95% CI [min]
IV, Random, 95% CI [min]
Glass 2008 396 78 19 378 78 19 19.0% 18.00 [-31.60, 67.60]
Morin 2005 419.59 60.62 58 401.76 75.35 65 81.0% 17.83 [-6.23, 41.89]
Total (95% CI) 77 84 100.0% 17.86 [-3.79, 39.51]
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%
Test for overall effect: Z = 1.62 (P = 0.11)
-50
-25
0
25
50
Favours Placebo
Favours 50 mg Diphenhydramine

Table S21 – Summary of Findings table for diphenhydramine 50 mg for the treatment of chronic insomnia

References: Glass 2008(A); Morin 2005(B)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 50 mg Diphenhydramine vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[5,7]	The mean sleep latency in the diphenhydramine group was **7.89 minutes lower** (17.40 lower to 1.62 higher)	52 (1 study)[A]
Sleep Latency (Subjective)	⊕⊕⊖⊖ **low**[1,2]	The mean sleep latency in the diphenhydramine groups was **2.47 minutes lower** (8.17 lower to 3.23 higher)	163 (2 studies)[A,B]
Total Sleep Time* (PSG)	⊕⊕⊖⊖ **low**[5,8]	The mean total sleep time in the diphenhydramine group was **12.37 minutes higher** (13.38 lower to 38.12 higher)	52 (1 study)[A]
Total Sleep Time (Subjective)	⊕⊕⊖⊖ **low**[1,2]	The mean total sleep time in the diphenhydramine groups was **17.86 minutes higher** (3.79 lower to 39.51 higher)	161 (2 studies)[A,B]
Quality of Sleep* (Subjective)	⊕⊕⊕⊖ **moderate**[5]	The mean quality of sleep in the diphenhydramine group was **0.1 points**[9] **higher** (0.45 lower to 0.65 higher)	38 (1 study)[A]
Sleep Efficiency (PSG)	⊕⊕⊖⊖ **low**[4,5]	The mean sleep efficiency in the diphenhydramine group was **2.59 percent higher** (3.25 lower to 8.43 higher)	52 (1 study)[B]
Sleep Efficiency (Subjective)	⊕⊕⊕⊖ **moderate**[5]	The mean sleep efficiency in the diphenhydramine group was **4.61 percent higher** (1.33 to 7.88 higher)	123 (1 study)[A]
Number of Awakenings (Subjective)	⊕⊕⊕⊖ **moderate**[3]	The mean number of awakenings in the diphenhydramine group was **0.3 awakenings lower** (1.03 lower to 0.43 higher)	38 (1 study)[A]

* Critical Outcome, used to determine Quality of Evidence
[1] SL and TST 95% Ci cross Clinical Significance
[2] 1 of 2 studies funded by industry
[3] 95% CI (-1.03, 0.43) crosses Clinical Significance (0.5 awakenings)
[4] 95% CI (-3.25, 8.43) crosses Clinical Significance (5%)
[5] Study funded by industry
[6] 95% CI (-0.45, 0.65) crosses zero standard mean difference
[7] 95% CI (-17.4, 1.62) crosses Clinical Significance (10 minutes)
[8] 95% CI (-13.38, 38.12) crosses Clinical Significance (20 minutes)
[9] 5-point scale (higher score indicates better sleep quality)

Melatonin - Meta-Analyses and Summary of Findings Tables

Figure S88 – Meta-analysis of data for subjectively-determined quality of sleep in response to melatonin 2 mg

Study or Subgroup	2 mg Melatonin Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Lemoine, 2007	26.7	21.5	44	17.5	16.6	43	33.5%	0.47 [0.05, 0.90]
Luthringer, 2009	42.1	14.8	20	33.8	17.3	20	27.3%	0.51 [-0.13, 1.14]
Wade, 2007	45.9	16	169	49.5	14.8	165	39.2%	-0.23 [-0.45, -0.02]
Total (95% CI)			233			228	100.0%	**0.21 [-0.36, 0.77]**

Heterogeneity: Tau² = 0.20; Chi² = 11.54, df = 2 (P = 0.003); I² = 83%
Test for overall effect: Z = 0.72 (P = 0.47)

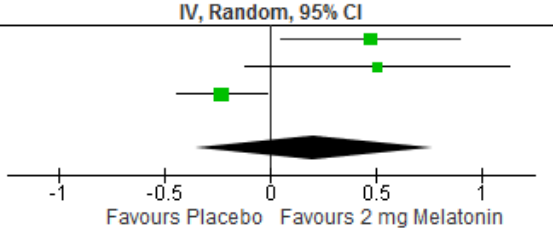


Table S22 – Summary of Findings table for melatonin 2 mg for the treatment of chronic insomnia

References: Lemoine 2007(A); Luthringer 2009(B); Wade 2007(C)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 2 mg Melatonin vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[3,4]	The mean sleep latency in the melatonin group was **8.9 minutes lower** (15.45 to 2.35 lower)	40 (1 study)[B]
Total Sleep Time (PSG)	⊕⊖⊖⊖ **very low**[3,6]	The mean total sleep time in the melatonin group was **2.2 minutes higher** (19.13 lower to 23.53 higher)	40 (1 study)[B]
Wake After Sleep Onset (PSG)	⊕⊕⊖⊖ **low**[3,5]	The mean wake after sleep onset in the melatonin group was **8.5 minutes higher** (11.75 lower to 28.75 higher)	40 (1 study)[B]
Quality of Sleep* (Subjective)	⊕⊖⊖⊖ **very low**[1,2,3]	The mean quality of sleep in the melatonin group was **0.21 standard deviations higher** (0.36 lower to 0.77 higher)	461 (3 studies)[A,B,C]
Number of Awakenings (PSG)	⊕⊖⊖⊖ **very low**[3,7]	The mean number of awakenings in the melatonin group was **1.4 awakenings higher** (4.59 lower to 7.39 higher)	40 (1 study)[B]

* Critical Outcome, used to determine Quality of Evidence
[1] Heterogeneity (I^2 = 83%) greater than allowance (75%)
[2] 95% CI (-0.36, 0.77) crosses zero standard mean difference
[3] All studies funded by industry
[4] 95% CI (-15.45, -2.35) crosses Clinical Significance (10 min)
[5] 95% CI (-11.75, 28.75) crosses Clinical Significance (20 min)
[6] 95% CI (-19.13, 23.53) crosses Clinical Significance (20 min)
[7] 95% CI (-4.59, 7.39) crosses Clinical Significance

L-tryptophan - Summary of Findings Table

Table S23 – Summary of Findings table for L-tryptophan 250 mg for the treatment of chronic insomnia

Reference: Hudson 2005

Outcomes	Quality of the evidence (GRADE)	Absolute Difference 250 mg Tryptophan vs Placebo	No of Participants (studies)
Wake After Sleep Onset* (Subjective)	⊕⊕⊕⊕ **high**	The mean wake after sleep onset in the Tryptophan groups was **9.70 minutes lower** (15.21 to 4.18 lower)	31 (1 study)
Total Sleep Time (Subjective)	⊕⊕⊕⊖ **moderate**[1]	The mean total sleep time in the Tryptophan groups was **20.00 minutes lower** (31.29 to 8.7 lower)	32 (1 study)
Quality of Sleep* (Subjective)	⊕⊕⊕⊕ **high**	The mean quality of sleep in the Tryptophan groups was **0.30 points**[2] **higher** (0.22 to 0.37 higher)	32 (1 study)
Sleep Efficiency (Subjective)	⊕⊕⊕⊕ **high**	The mean sleep efficiency in the Tryptophan groups was **2.20 percent lower** (4.27 to 0.12 lower)	32 (1 study)

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (8.7, 31.29) crosses Clinical Significance (30 min)
[2] 3-point scale (Sleep Quality index: 1=low, 3=high)

Valerian - Summary of Findings Table

Table S24 – Summary of Findings table for valerian for the treatment of chronic insomnia

References: Morin 2005(A)

Outcomes	Quality of the evidence (GRADE)	Absolute Difference Valerian-hops vs Placebo	No of Participants (studies)
Sleep Latency* (PSG)	⊕⊕⊖⊖ **low**[1,2]	The mean sleep latency in the Valerian-hops groups was **9.29 minutes lower** (18.3 to 0.27 lower)	48 (1 study) [A]
Sleep Latency (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean sleep latency in the Valerian-hops groups was **3.77 minutes higher** (4.47 lower to 12.01 higher)	124 (1 study) [A]
Total Sleep Time (PSG)	⊕⊖⊖⊖ **very low**[2,3]	The mean total sleep time in the Valerian-hops groups was **10.96 minutes higher** (21.67 lower to 43.59 higher)	48 (1 study) [A]
Total Sleep Time (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean total sleep time in the Valerian-hops groups was **3.12 minutes higher** (22.08 lower to 28.32 higher)	123 (1 study) [A]
Sleep Efficiency (PSG)	⊕⊖⊖⊖ **very low**[2,4]	The mean sleep efficiency in the Valerian-hops groups was **0.96 percent higher** (5.02 lower to 6.94 higher)	48 (1 study) [A]
Sleep Efficiency (Subjective)	⊕⊕⊕⊖ **moderate**[2]	The mean sleep efficiency in the Valerian-hops groups was **1.85 percent higher** (1.9 lower to 5.6 higher)	123 (1 study) [A]

* Critical Outcome, used to determine Quality of Evidence
[1] 95% CI (-18.3, -0.27) crosses Clinical Significance (10 min)
[2] Study funded by industry
[3] 95% CI (-21.67, 43.59) crosses Clinical Significance (20 min)
[4] 95% CI (-5.02, 6.94) crosses Clinical Significance (5%)